

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2005, 13:22:36 ; Search time 118 Seconds
(without alignments)
22.943 Million cell updates/sec

Title: 09730329-60ED

Perfect score: 29

Sequence: 1 VMXVAEF 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 300 summaries

Database : A_Geneseq_16Dec04:*

1: Geneseqp1980s:*

2: Geneseqp1990s:*

3: Geneseqp2000s:*

4: Geneseqp2001s:*

5: Geneseqp2002s:*

6: Geneseqp2003as:*

7: Geneseqp2003bs:*

8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	27	93.1	7	3 AAB07871	Aab07871 A beta-se
2	27	93.1	8	3 AAB07872	Aab07872 A beta-se
3	27	93.1	68	8 ADJ84527	Adj84527 Human T2R
4	27	93.1	450	7 ADF06975	Adf06975 Bacterial
5	27	93.1	457	8 ADN46338	Adn46338 Thermococ
6	27	93.1	802	5 ABB93074	Abb93074 Herbicida
7	26	89.7	223	6 ABU35492	Abu35492 Protein e
8	26	89.7	224	6 ABU36167	Abu36167 Protein e
9	26	89.7	304	8 ADN47550	Adn47550 Thermococ
10	26	89.7	821	2 AAY34479	Aay34479 Porphorym
11	26	89.7	869	2 AAY34354	Aay34354 Porphorym
12	25	86.2	194	7 ADH85915	Adh85915 Enterococ
13	25	86.2	406	6 ABU29329	Abu29329 Protein e
14	25	86.2	583	8 ADN48186	Adn48186 Thermococ
15	24	82.8	8	8 ADS09499	Ads09499 Beta-recr
16	24	82.8	64	5 ABP03038	Abp03038 Human ORF
17	24	82.8	75	4 AAU56778	Aau56778 Propionib
18	24	82.8	75	6 ABM53297	Abm53297 Propionib
19	24	82.8	97	6 ABP75586	Abp75586 Human sec
20	24	82.8	132	4 AAM41985	Aam41985 Human pol
21	24	82.8	132	4 AAM41986	Aam41986 Human pol
22	24	82.8	132	4 AAM41984	Aam41984 Human pol
23	24	82.8	145	5 ABG98409	Abg98409 Haloarcul
24	24	82.8	191	6 ABU24614	Abu24614 Protein e
25	24	82.8	195	6 ABU23979	Abu23979 Protein e

26	24	82.8	195	8 ADJ48536	Adj48536 Oil-assoc
27	24	82.8	225	2 AAY85820	Aay85820 S. pneumo
28	24	82.8	236	4 AAU50802	Aau50802 Propionib
29	24	82.8	236	6 ABM47321	Abm47321 Propionib
30	24	82.8	238	6 ABU26266	Abu26266 Protein e
31	24	82.8	254	2 AAY36930	Aay36930 Amino aci
32	24	82.8	286	4 ABG18334	Abg18334 Novel hum
33	24	82.8	332	5 ABB81185	Abb81185 VAB-1 pro
34	24	82.8	335	4 AAY37753	Aay37753 Streptoco
35	24	82.8	335	6 ABU01294	Abu01294 S. pneumo
36	24	82.8	335	6 ABP81576	Abp81576 Streptoco
37	24	82.8	335	6 ABU45990	Abu45990 Protein e
38	24	82.8	335	8 ADK48584	Adk48584 Streptoco
39	24	82.8	361	3 AAG20945	Aag20945 Arabidops
40	24	82.8	361	3 AAG45883	Aag45883 Arabidops
41	24	82.8	361	3 AAG24458	Aag24458 Arabidops
42	24	82.8	367	6 ABU23606	Abu23606 Protein e
43	24	82.8	373	4 AAM40200	Aam40200 Human pol
44	24	82.8	378	3 AAG20944	Aag20944 Arabidops
45	24	82.8	378	3 AAG24457	Aag24457 Arabidops
46	24	82.8	378	3 AAG45882	Aag45882 Arabidops
47	24	82.8	384	3 AAY71056	Aay71056 Human mem
48	24	82.8	384	4 AAM40198	Aam40198 Human pol
49	24	82.8	384	4 AAM40199	Aam40199 Human pol
50	24	82.8	409	5 AAE30494	Aae30494 Fruit fly
51	24	82.8	414	3 AAG45881	Aag45881 Arabidops
52	24	82.8	414	3 AAG24456	Aag24456 Arabidops
53	24	82.8	414	3 AAG20943	Aag20943 Arabidops
54	24	82.8	457	8 ADR09407	Adr09407 Human pro
55	24	82.8	584	4 AAB96296	Aab96296 Putative
56	24	82.8	807	6 ABU18136	Abu18136 Protein e
57	24	82.8	817	4 ABB70027	Abb70027 Drosophil
58	24	82.8	902	6 ABU18144	Abu18144 Protein e
59	24	82.8	919	6 ABU42813	Abu42813 Protein e
60	24	82.8	919	8 ADN22714	Adn22714 Bacterial
61	24	82.8	926	6 ADB11880	Adb11880 Alloiococ
62	24	82.8	927	5 ABP39986	Abp39986 Staphyloc
63	24	82.8	927	8 ADS06396	Ads06396 Staphyloc
64	24	82.8	938	6 ADB11878	Adb11878 Alloiococ
65	24	82.8	940	6 ADB11876	Adb11876 Alloiococ
66	24	82.8	962	2 AAR58701	Aar58701 Aminopept
67	24	82.8	972	2 AAR51280	Aar51280 Helminth
68	24	82.8	1084	5 ABP52161	Abp52161 Human 670
69	24	82.8	1084	5 ABP52175	Abp52175 Human 670
70	24	82.8	1084	7 AAD37496	Adi37496 Human tra
71	24	82.8	1084	8 ADI27975	Adi27975 Human 670
72	24	82.8	1084	8 ADN33134	Adn33134 Human tra
73	24	82.8	1095	5 ABP52162	Abp52162 Mouse pot
74	24	82.8	1095	5 ABP52170	Abp52170 Mouse pot
75	24	82.8	1095	6 AAE32082	Aae32082 Human TRI
76	24	82.8	1095	7 AAD37509	Adi37509 Mouse pho
77	24	82.8	1095	7 AAD37499	Adi37499 Human tra
78	24	82.8	1095	8 ADI27978	Adi27978 Human 670
79	24	82.8	1095	8 ADI27988	Adi27988 Murine pr
80	24	82.8	1176	7 ABO62362	Abu62362 Klebsiell
81	24	82.8	1225	6 ABU28134	Abu28134 Protein e
82	23	79.3	36	2 AAW27782	Aaw27782 UDP-N-ace
83	23	79.3	54	2 AAW79151	Aaw79151 Receptor
84	23	79.3	54	2 AAW81408	Aaw81408 Receptor
85	23	79.3	54	6 ABU83035	Abu83035 Mouse rec
86	23	79.3	54	6 ABU83019	Abu83019 Rat recep
87	23	79.3	54	6 ABU83034	Abu83034 Receptor
88	23	79.3	56	5 AAW78556	Aaw78556 Residues
89	23	79.3	68	5 ABP31047	Abp31047 Human ORF
90	23	79.3	84	4 AAU41413	Aau41413 Propionib
91	23	79.3	84	6 ABM37932	Abm37932 Propionib
92	23	79.3	90	4 AAB87752	Abb87752 Human T2R
93	23	79.3	90	4 ADJ84523	Adj84523 Human T2R
94	23	79.3	90	8 ADR29101	Adr29101 Taste rec
95	23	79.3	112	7 ADC01644	Adc01644 Enterohae
96	23	79.3	123	7 ADM25825	Adm25825 Hyperther
97	23	79.3	132	7 ADH85757	Adh85757 Enterococ
98	23	79.3	143	3 AAB10816	Aab10816 Archaeogl

99	23	79.3	150	4	AU33403	Aau33403 Enterococ	172	23	79.3	457	6	ABP76802	Abp76802 N. gonorr
100	23	79.3	162	5	ABBS5092	Abbs5092 Lactococc	173	23	79.3	457	6	ABP80055	Abp80055 N. gonorr
101	23	79.3	162	8	ADS29448	Ads29448 Bacterial	174	23	79.3	483	7	ABO82504	AbO82504 Pseudom
102	23	79.3	189	4	ABBS2817	Abbs2817 Escherich	175	23	79.3	484	8	ADN23529	Adn23529 Bacterial
103	23	79.3	191	6	ABU49757	Abu49757 Protein e	176	23	79.3	526	5	AEE24993	Aee24993 Human fib
104	23	79.3	205	7	ADM25883	Adm25883 Hyperther	177	23	79.3	595	6	ABM68180	Abm68180 Photorhab
105	23	79.3	209	6	ABU44637	Abu44637 Protein e	178	23	79.3	613	6	ADA36704	Ada36704 Acinetoba
106	23	79.3	237	4	AAG90704	Aag90704 C glutami	179	23	79.3	637	8	ABM83556	Abm83556 Human dia
107	23	79.3	244	7	ADB64338	Adb64338 Human pro	180	23	79.3	648	2	AAR08287	Aar08287 NSI prote
108	23	79.3	245	8	ADR15793	Adr15793 Kinase 1f	181	23	79.3	650	7	ADE58201	Ades58201 Rat Prote
109	23	79.3	246	4	ABBS8040	Abbs8040 Drosophil	182	23	79.3	650	7	ADE58205	Ades58205 Rat Prote
110	23	79.3	247	7	ADE86966	Ade86966 Human pan	183	23	79.3	650	7	ADE58197	Ades58197 Rat Prote
111	23	79.3	249	4	AAE13127	Aae13127 Sheep SLS	184	23	79.3	650	7	ADE58193	Ades58193 Rat Prote
112	23	79.3	249	4	AAE13136	Aae13136 Sheep SLS	185	23	79.3	650	7	ADSO3534	Adeo3534 Rat fibro
113	23	79.3	254	7	ADG63616	Adg63616 Disease t	186	23	79.3	652	2	AAR15269	Aar15269 Clone pfb
114	23	79.3	254	8	ADG65401	Adg65401 S. cerevi	187	23	79.3	654	7	ADD47011	Add47011 Human pro
115	23	79.3	274	5	ABP52383	Abp52383 Human JAK	188	23	79.3	660	8	ABM83557	Abm83557 Human dia
116	23	79.3	280	4	AAG90751	Aag90751 C glutami	189	23	79.3	672	8	ADO26840	Ado26840 Human rec
117	23	79.3	287	2	AAU03631	Aau03631 Amino aci	190	23	79.3	679	2	AAR41517	Aar41517 K-sam. 3/
118	23	79.3	290	8	ADR88390	Adr88390 FGFR1 tyr	191	23	79.3	682	7	ADH63162	Adh63162 Human FGF
119	23	79.3	293	6	ABU25031	Abu25031 Protein e	192	23	79.3	687	2	AAU06456	Aay06456 ZNF198-FG
120	23	79.3	295	5	ABB81187	Abb81187 Human FGF	193	23	79.3	694	5	ABG66739	Abg66739 Human nov
121	23	79.3	299	4	ABBG63276	Abbg63276 Drosophil	194	23	79.3	694	5	ABG66739	Abg66739 Human nov
122	23	79.3	299	8	ADR15783	Adr15783 Kinase la	195	23	79.3	702	5	ABG79680	Abg79680 Tumour in
123	23	79.3	299	8	ADR15783	Adr15783 Kinase la	196	23	79.3	705	7	ADH63163	Adh63163 Human FGF
124	23	79.3	301	5	ABP52381	Abp52381 Human JAK	197	23	79.3	726	2	AAR14280	Aar14280 Murine KG
125	23	79.3	309	8	ADS22085	Ads22085 Bacterial	198	23	79.3	726	2	AAR10933	Aar10933 KGF recep
126	23	79.3	310	2	AAW55999	Aaw55999 Protein t	199	23	79.3	729	2	AAW37804	Aaw37804 Fibroblas
127	23	79.3	313	3	AAU59053	Aau59053 Murine F1	200	23	79.3	729	2	AAU23631	Aau23631 Bek-like
128	23	79.3	314	4	AAU87775	Aau87775 Human T2R	201	23	79.3	731	2	AAU10649	Aar10649 Chicken b
129	23	79.3	314	4	AAU87754	Aau87754 Human T2R	202	23	79.3	731	2	AAU21686	Aar21686 Human bFG
130	23	79.3	314	5	ABP95931	Abp95931 Human GPC	203	23	79.3	731	2	AAU23633	Aay23633 Newt kera
131	23	79.3	314	7	ADC87025	Adc87025 Human GPC	204	23	79.3	731	8	ABM83558	Abm83558 Human dia
132	23	79.3	314	8	ADR29132	Adr29132 Taste rec	205	23	79.3	735	8	ADO42157	Ado42157 Human NOV
133	23	79.3	314	8	ADR29105	Adr29105 Taste rec	206	23	79.3	749	8	ABM83559	Abm83559 Human dia
134	23	79.3	314	8	ADM33333	Adm33333 Human bit	207	23	79.3	750	2	AAU37806	Aaw37806 Keratinoc
135	23	79.3	315	2	AAW56001	Aaw56001 Protein t	208	23	79.3	762	6	ABR58628	AbR58628 Human can
136	23	79.3	335	7	ADM26176	Adm26176 Hyperther	209	23	79.3	762	8	ABM80714	Abm80714 Tumour-ae
137	23	79.3	344	5	ABP40722	Abp40722 Staphyloc	210	23	79.3	764	3	ABM58376	AbM58376 Lung canc
138	23	79.3	344	8	ADS08018	Ads08018 Staphyloc	211	23	79.3	768	7	ADH63160	Adh63160 Human FGF
139	23	79.3	345	8	ADO60028	Ado60028 CRH signa	212	23	79.3	769	2	AAR15268	Aar15268 Clone pfb
140	23	79.3	345	8	ADO44548	Ado44548 Fibroblas	213	23	79.3	771	7	ADH63158	Adh63158 Human FGF
141	23	79.3	351	2	AAW56000	Aaw56000 Fibroblas	214	23	79.3	773	4	AAU35737	Aau35737 Helicobac
142	23	79.3	378	2	AAW37805	Aaw37805 Fibroblas	215	23	79.3	773	6	ABU30798	Abu30798 Protein e
143	23	79.3	378	2	AAU23632	Aau23632 Newt fibr	216	23	79.3	783	8	ABM83106	Abm83106 Human dia
144	23	79.3	383	6	ABU15985	Abu15985 Protein e	217	23	79.3	789	8	ADH181626	Adh181626 Human pro
145	23	79.3	387	6	AAU33637	Aau33637 Pseudom	218	23	79.3	789	8	ADL57085	Adl57085 Human NOV
146	23	79.3	387	6	ABU15630	Abu15630 Protein e	219	23	79.3	789	8	ADL57093	Adl57093 Human NOV
147	23	79.3	388	2	AAW37803	Aaw37803 Truncated	220	23	79.3	789	8	ADL57087	Adl57087 Human NOV
148	23	79.3	388	2	AAU23630	Aau23630 Truncated	221	23	79.3	789	8	ADL57063	Adl57063 Human NOV
149	23	79.3	393	4	AAW96253	Aaw96253 Putative	222	23	79.3	789	8	ADL57065	Adl57065 Human NOV
150	23	79.3	393	8	ADR28234	Adr28234 Pyrococcu	223	23	79.3	789	8	ADL57071	Adl57071 Human NOV
151	23	79.3	393	8	ADS43183	Ads43183 Bacterial	224	23	79.3	789	8	ADL57073	Adl57073 Human NOV
152	23	79.3	394	7	ABM85572	Abm85572 Mouse pro	225	23	79.3	789	8	ADL57083	Adl57083 Human NOV
153	23	79.3	397	7	ABO79676	AbO79676 Pseudom	226	23	79.3	789	8	ADL57077	Adl57077 Human NOV
154	23	79.3	399	5	ABBA7692	Abba7692 Listeria	227	23	79.3	789	8	ADL57089	Adl57089 Human NOV
155	23	79.3	421	2	AAW96155	Aaw96155 Staphyloc	228	23	79.3	789	8	ADL57061	Adl57061 Human NOV
156	23	79.3	421	3	AAU70130	Aau70130 Staphyloc	229	23	79.3	789	8	ADL57079	Adl57079 Human NOV
157	23	79.3	421	3	AAU70137	Aau70137 Staphyloc	230	23	79.3	789	8	ADL57057	Adl57057 Human NOV
158	23	79.3	421	4	AAU36795	Aau36795 Staphyloc	231	23	79.3	789	8	ADL57075	Adl57075 Human NOV
159	23	79.3	421	4	AAU37227	Aau37227 Staphyloc	232	23	79.3	789	8	ADL57059	Adl57059 Human NOV
160	23	79.3	421	4	AAU33807	Aau33807 Staphyloc	233	23	79.3	789	8	ADL57069	Adl57069 Human NOV
161	23	79.3	421	4	AAG82362	Aag82362 S. epider	234	23	79.3	789	8	ADL57067	Adl57067 Human NOV
162	23	79.3	421	5	AAE25686	Aae25686 S. aureus	235	23	79.3	789	8	ADL57081	Adl57081 Human NOV
163	23	79.3	421	5	AAE25679	Aae25679 S. aureus	236	23	79.3	789	8	ADL57091	Adl57091 Human NOV
164	23	79.3	421	6	ABU43299	Abu43299 Protein e	237	23	79.3	789	8	ADL57095	Adl57095 Human NOV
165	23	79.3	421	6	ABM71834	Abm71834 Staphyloc	238	23	79.3	792	4	ABM83854	Abm83854 Human dia
166	23	79.3	421	7	ADL46298	Adl46298 UDP-N-ace	239	23	79.3	801	4	AAE21977	Aae21977 Mouse fib
167	23	79.3	421	7	ADL46296	Adl46296 UDP-N-ace	240	23	79.3	802	5	AAE16588	Aae16588 Human fib
168	23	79.3	426	6	ABU25359	Abu25359 Protein e	241	23	79.3	802	6	ABR58627	AbR58627 Human can
169	23	79.3	426	2	AAR26278	Aar26278 Tyrosine	242	23	79.3	802	7	ADB31984	Adb31984 Human fib
170	23	79.3	431	5	ABP39412	Abp39412 Staphyloc	243	23	79.3	802	7	ADE58199	Ades58199 Human pro
171	23	79.3	431	8	ADS04616	Ads04616 Staphyloc	244	23	79.3	802	7	ADE58203	Ades58203 Human pro

245 23 79.3 802 7 ADE58195
246 23 79.3 802 7 ADE58207
247 23 79.3 802 7 ADE38383
248 23 79.3 802 7 ADF45060
249 23 79.3 802 8 ADO49093
250 23 79.3 802 8 ADBM80713
251 23 79.3 804 8 ABM83110
252 23 79.3 806 4 AAE21976
253 23 79.3 806 6 ABJ38646
254 23 79.3 806 6 ADA89050
255 23 79.3 806 6 AAE36462
256 23 79.3 806 7 ADF45059
257 23 79.3 806 7 ADG74302
258 23 79.3 806 7 ADK51038
259 23 79.3 808 7 ADK51040
260 23 79.3 816 2 AAR39647
261 23 79.3 817 2 ADH63159
262 23 79.3 819 2 AAR10648
263 23 79.3 819 2 ADH63157
264 23 79.3 820 2 AAR47233
265 23 79.3 820 2 AAW63844
266 23 79.3 820 2 AAY06458
267 23 79.3 820 4 AAB84383
268 23 79.3 820 8 ADO49087
269 23 79.3 820 8 ADS18125
270 23 79.3 820 8 ADS18122
271 23 79.3 820 8 ADS18132
272 23 79.3 821 2 AAR21080
273 23 79.3 821 7 ADE57113
274 23 79.3 821 7 ADD45446
275 23 79.3 821 7 ADF45058
276 23 79.3 821 7 ADH63155
277 23 79.3 821 7 ADH63156
278 23 79.3 821 8 ADO49089
279 23 79.3 822 2 AAR13549
280 23 79.3 822 2 AAR20750
281 23 79.3 822 2 AAR26337
282 23 79.3 822 4 AAU04694
283 23 79.3 822 6 AAE34488
284 23 79.3 822 7 ABR56164
285 23 79.3 822 7 ADE61698
286 23 79.3 822 7 ADE55803
287 23 79.3 822 7 ADE61700
288 23 79.3 822 7 ADF45057
289 23 79.3 822 7 ADH63164
290 23 79.3 822 7 ADP65260
291 23 79.3 822 7 ADP65261
292 23 79.3 822 8 ADL61330
293 23 79.3 822 8 ADN03976
294 23 79.3 822 8 ADO42159
295 23 79.3 822 8 ADO17658
296 23 79.3 822 8 ADQ80367
297 23 79.3 822 8 ADP24447
298 23 79.3 824 8 ADO42155
299 23 79.3 830 8 ABM83108
300 23 79.3 831 8 ABM83107

ALIGNMENTS

RESULT 1

AAB07871
ID AAB07871 standard; peptide; 7 AA.

XX AC AAB07871;

XX DT 14-NOV-2000 (first entry)

XX DE A beta-secretase inhibitor peptide.

XX KW Beta-secretase; beta-amyloid precursor protein; beta-amyloid peptide;
KW amyloid plaque component; Alzheimer's disease; amyloidogenic disease;

KW inhibitor.
XX Synthetic.
XX OS
FH Key Location/Qualifiers
FT Modified-site 3 /note= "hydroxyethylene"
FT XX
PN WO200047618-A2.
XX PD 17-AUG-2000.
XX XX
PF 10-FEB-2000; 2000WO-US003819.
XX XX
PR 10-FEB-1999; 99US-0119571P.
PR 15-JUN-1999; 99US-0139172P.
XX (ELAN-) ELAN PHARM INC.
XX PA Anderson JP, Basi G, Doane MT, Frigon N, John V, Power M;
PI Sinha S, Tatsuno G, Tung J, Wang S, Mcconlogue L;
XX WPI; 2000-533011/48.
DR Purified beta-secretase protein used in assays to discover inhibitors
XX which can be used for the treatment of amyloidogenic diseases e.g.
XX PT Alzheimer's disease.
XX XX
PS Disclosure; Page 12; 121pp; English.
XX XX
CC The specification describes a beta-secretase enzyme. The enzyme cleaves
CC beta-amyloid precursor protein to produce beta-amyloid peptide. This
CC enzyme is therefore implicated in the production of amyloid plaque
CC components which accumulate in the brains of individuals afflicted with
CC Alzheimer's disease. Inhibitors of beta-secretase are administered to a
CC mammalian subject e.g. with Alzheimer's disease or Alzheimer's disease-
CC like pathology to test if they maintain or improve cognitive ability or
CC reduce the plaque burden. The compounds are used for the treatment of
CC amyloidogenic diseases e.g. Alzheimer's disease. The present sequence
CC represents an inhibitor of beta-secretase enzyme
XX Sequence 7 AA;
SQ
Query Match 93.1%; Score 27; DB 3; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.8e+06;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VMXVAEF 7
Db 1 VMXVAEF 7
RESULT 2
AAB07872
ID AAB07872 standard; peptide; 8 AA.
XX AC AAB07872;
XX XX
DT 14-NOV-2000 (first entry)
XX DE A beta-secretase inhibitor peptide.
XX KW Beta-secretase; beta-amyloid precursor protein; beta-amyloid peptide;
KW amyloid plaque component; Alzheimer's disease; amyloidogenic disease;
XX KW inhibitor.
XX OS Synthetic.
XX XX
FH Key Location/Qualifiers
FT Modified-site 4 /note= "hydroxyethylene"
FT XX
PN WO200047618-A2.

XX 17-AUG-2000.
 XX 10-FEB-2000; 2000WO-US003819.
 XX 10-FEB-1999; 99US-0119571P.
 XX 15-JUN-1999; 99US-0139172P.
 XX (ELAN-) ELAN PHARM INC.
 XX Anderson JP, Basi G, Doane MT, Frigon N, John V, Power M;
 PI Sinha S, Tatsuno G, Tung J, Wang S, Mcconlogue L;
 XX WPI; 2000-533011/48.
 XX Purified beta-secretase protein used in assays to discover inhibitors
 PT which can be used for the treatment of amyloidogenic diseases e.g.
 PT Alzheimer's disease.
 XX Alzheimer's disease. Inhibitors of beta-secretase are administered to a
 XX mammalian subject e.g. with Alzheimer's disease or Alzheimer's disease-
 XX like pathology to test if they maintain or improve cognitive ability or
 XX reduce the plaque burden. The compounds are used for the treatment of
 XX amyloidogenic diseases e.g. Alzheimer's disease. The present sequence
 XX represents an inhibitor of beta-secretase enzyme
 XX Sequence 8 AA;
 SQ Query Match 93.1%; Score 27; DB 3; Length 8;
 Best Local Similarity 100.0%; Pred. No. 1.8e+06;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 VMXVAEF 7
 Db 2 VMXVAEF 8
 RESULT 3
 ADJ84527
 ID ADJ84527 standard; protein; 68 AA.
 XX ADJ84527;
 XX 20-MAY-2004 (first entry)
 XX Human T2R G-protein coupled receptor seq id 80.
 XX taste transduction; G-protein coupled receptor; T2R; taste signaling;
 KW CAMP level; cGMP level; IP3 level; Ca(2+) level; electrical activity;
 KW human; G-protein coupled receptor; receptor.
 XX Homo sapiens.
 OS US2004038312-A1.
 XX 26-FEB-2004.
 XX 10-FEB-2003; 2003US-00364861.
 XX 10-SEP-1999; 99US-00393634.
 XX (ZUKER/) ZUKER C S.
 PA (ADLER/) ADLER J E.
 PA (HOON/) HOON M.
 PA (RYBA/) RYBA N.
 PA (MUEL/) MUELLER K.
 XX

PI Zuker CS, Adler JE, Hoon M, Ryba N, Mueller K;
 XX WPI; 2004-203221/19.
 DR N-PSDB; ADJ84528.
 XX Novel isolated taste transduction G-protein coupled receptor e.g., T2R
 PT useful for identifying compound that modulates taste signaling in taste
 PT cells.
 XX Claim 22; SEQ ID NO 80; 121pp; English.
 XX The invention describes an isolated taste transduction G-protein coupled
 CC receptor (I) e.g., T2R which is expressed in a taste cell, comprises
 CC greater than 60% sequence identity to a fully defined sequence of 335
 CC (S1), 333 (S2), 299 (S3), 310 (S4), 224 (S5), 77 (S6), 209 (S7), 266
 CC (S8), 300 (S9), 300 (S10), 155 (S11), 173 (S12), 94 (S13), 115 (S14), 68
 CC (S15), 126 (S16) and 180 (S17) amino acids as given in the specification.
 CC (I) is useful for identifying a compound that modulates taste signaling
 CC in taste cells which involves contacting the compound with (I) and
 CC determining the functional effect of the compound (I). The functional
 CC effect is determined by measuring changes in intracellular cAMP, cGMP,
 CC IP3, or Ca (2+). The functional effect is a chemical or physical effect.
 CC The functional effect is determined by measuring binding of the compound
 CC to an extracellular domain of (I). The functional effect is determined by
 CC measuring binding of radiolabeled GTP to (I). (I) is recombinant and is
 CC from rat, mouse, or human. (I) is expressed in a cell or cell membrane.
 CC The functional effect is measured by determining changes in the
 CC electrical activity of cells (e.g., eukaryotic cell) expressing (I). (I)
 CC is useful for identifying compound that modulates taste signaling in
 CC taste cells which involves contacting a compound with (I). The
 CC extracellular domain (II) of (I) is useful for identifying a compound
 CC that modulates taste signaling in taste cells which involves contacting
 CC the compound with the polynucleotide encoding (I) and determining the
 CC functional effect of the compound upon the extracellular domain. This
 CC sequence encodes a human T2R G-protein coupled receptor.
 XX Sequence 68 AA;
 SQ Query Match 93.1%; Score 27; DB 8; Length 68;
 Best Local Similarity 71.4%; Pred. No. 23;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 VMXVAEF 7
 Db 11 IMAVAEF 17
 RESULT 4
 ADF06975
 ID ADF06975 standard; protein; 450 AA.
 XX ADF06975;
 XX 12-FEB-2004 (first entry)
 XX Bacterial polypeptide #3088.
 DE Proteus mirabilis infection; bacterial infection; antibacterial;
 XX immunostimulant.
 KW Proteus mirabilis.
 XX US6605709-B1.
 PN 12-AUG-2003.
 XX 05-APR-2000; 2000US-00543681.
 XX 09-APR-1999; 99US-0128706P.
 PR (GENO-) GENOME THERAPEUTICS CORP.
 XX Breton GL;
 PI

XX WPI; 2003-895291/82.
 DR N-PSDB; ADF02803.
 XX
 PT New Proteus mirabilis polypeptides and polynucleotides, useful as
 PT reagents for diagnosis of bacterial disease, as components of
 PT antibacterial vaccines, as targets for antibacterial drugs, or as
 PT biocontrol agents for plants.
 XX
 PS Disclosure; SEQ ID NO 7260; 870pp; English.
 XX
 CC The invention relates to new Proteus mirabilis polypeptides and
 CC polynucleotides. The invention also relates to antibodies against the
 CC polypeptides, methods for producing the polypeptides, a method of
 CC generating vaccines for immunising an individual against P. mirabilis,
 CC method for evaluating a compound for the ability to bind a P. mirabilis
 CC polypeptide and a method for screening test compounds for anti-bacterial
 CC activity. The polypeptides and polynucleotides are useful as molecular
 CC targets for diagnosing, preventing and treating pathological conditions
 CC resulting from bacterial infection, as reagents for diagnosis of
 CC bacterial diseases, as components of antibacterial vaccines, as targets
 CC for antibacterial drugs or as bio-control agents for plants. This
 CC sequence represents a Proteus mirabilis polypeptide of the invention.
 XX
 SQ Sequence 450 AA;

Query Match 93.1%; Score 27; DB 7; Length 450;
 Best Local Similarity 85.7%; Pred. No. 1.8e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 VMXVAEF 7
 |||||
 Db 286 VMVVAEF 292

RESULT 5
 ADN46338
 ID ADN46338 standard; protein; 457 AA.
 AC ADN46338;
 XX
 DT 01-JUL-2004 (first entry)
 XX
 DE Thermococcus kodakaraensis KOD1 protein sequence SeqID216.
 XX
 KW gene disruption; gene targeting; marker gene; transformation;
 KW homologous recombination; hyperthermostable archaeobacterium; KOD1;
 KW gene structure; gene function; enzyme activity; medicine;
 KW forensic science; food; drug inspection; molecular biology; immunology.
 XX
 OS Thermococcus kodakaraensis.
 XX
 PN WO2004022736-A1.
 XX
 PD 18-MAR-2004.
 XX
 PF 29-AUG-2003; 2003WO-IB003597.
 XX
 PR 30-AUG-2002; 2002JP-00319011.
 XX
 PA (NISC-) JAPAN SCI & TECHNOLOGY CORP.
 XX
 PI Imanaka T, Atomi H;
 XX
 DR WPI; 2004-257583/24.
 XX
 CC Method for disrupting targeted gene in genome of organism particularly
 CC thermostable bacterium and with genome chips for analysis, applicable in
 CC studying gene structure and functions.
 XX
 PS Claim 9; SEQ ID NO 216; 598pp; Japanese.
 XX
 CC This invention relates to a novel method for targeting disruption of an

CC arbitrary gene in a genome of an organism which comprises providing the
 CC whole sequential data of the genome of such organism, selecting at least
 CC 1 arbitrary region in the sequence, providing a vector that contains a
 CC sequence homologous with the selected region and a marker gene,
 CC transformation, and homologous recombination. The genome is preferably
 CC the genome of a hyperthermostable archaeobacterium, particularly
 CC Thermococcus kodakaraensis KOD1. The method is for targeting the
 CC disruption of a gene in the genome of an organism, which is applicable in
 CC studying gene structure and functions as well as enzyme activities of
 CC encoded proteins and useful in medicine, forensic science, food or drug
 CC inspection, molecular biology and immunology. With this method, the
 CC disruption of a gene at an arbitrary position in a genome can be achieved
 CC efficiently and reliably. The present sequence is that of a protein
 CC encoded by the genome of Thermococcus kodakaraensis which was derived
 CC using the method of the invention. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 457 AA;

Query Match 93.1%; Score 27; DB 8; Length 457;
 Best Local Similarity 85.7%; Pred. No. 1.8e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 VMXVAEF 7
 |||||
 Db 1 VMEVVAEF 7

RESULT 6
 ABB93074
 ID ABB93074 standard; protein; 802 AA.
 AC ABB93074;
 XX
 DT 31-MAY-2002 (first entry)
 XX
 DE Herbicidally active polypeptide SEQ ID NO 2285.
 XX
 KW Herbicidal; plant; agriculture; herbicide.
 XX
 OS Arabidopsis thaliana.
 XX
 PN WO200210210-A2.
 XX
 PD 07-FEB-2002.
 XX
 PF 28-AUG-2001; 2001WO-EP009892.
 XX
 PR 28-AUG-2001; 2001WO-EP009892.
 XX
 PA (FARB) BAYER AG.
 XX
 PI Tietjen K, Weidler M;
 XX
 DR WPI; 2002-269010/31.
 XX
 PT Identifying plant target proteins for herbicidally active compounds,
 PT comprising aligning and comparing nucleic acid or amino acid sequences
 PT from plant with nucleic acid or amino acid sequences from non-plant
 PT organisms.
 XX
 PS Claim 5; SEQ ID NO 2285; 261pp + Sequence Listing; English.
 XX
 CC The invention relates to identifying target proteins (ABB90790-ABB94016)
 CC for herbicidally active compounds, comprising aligning and comparing
 CC nucleic acid or amino acid sequences from plant with nucleic acid or
 CC amino acid sequences from non-plant organisms using suitable search
 CC parameters, where plant sequences having an E-value greater by a factor
 CC of 3 than the E-value of most similar non-plant sequences are selected.
 CC The polypeptides or nucleic acids encoding them are useful for
 CC identifying modulators. The identified modulators are useful as

CC herbicides
XX Sequence 802 AA;
SQ

Query Match 93.1%; Score 27; DB 5; Length 802;
Best Local Similarity 71.4%; Pred. No. 3.4e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VMXVAEF 7
Db 629 VMSTAEF 635
|| :|||

RESULT 7
ID ABU35492 standard; protein; 223 AA.
XX AC ABU35492;
XX 19-JUN-2003 (first entry)
XX Protein encoded by Prokaryotic essential gene #21019.
XX Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX Mycoplasma genitalium.
XX WO200277183-A2.
XX 03-OCT-2002.
XX 21-MAR-2002; 2002WO-US009107.
XX 21-MAR-2001; 2001US-00815242.
XX 06-SEP-2001; 2001US-00948993.
XX 25-OCT-2001; 2001US-0342923P.
XX 08-FEB-2002; 2002US-00072851.
XX 06-MAR-2002; 2002US-0362699P.
XX (ELIT-) ELITRA PHARM INC.
XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
XX Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX WPI; 2003-029926/02.
XX N-PSDB; ACA39362.
XX New antisense nucleic acids, useful for identifying proteins or screening
XX for homologous nucleic acids required for cellular proliferation to
XX isolate candidate molecules for rational drug discovery programs.
XX Claim 25; SEQ ID NO 63416; 1766pp; English.
XX The invention relates to an isolated nucleic acid comprising any one of
XX the 6213 antisense sequences given in the specification where expression
XX of the nucleic acid inhibits proliferation of a cell. Also included are:
XX (1) a vector comprising a promoter operably linked to the nucleic acid
XX encoding a polypeptide whose expression is inhibited by the antisense
XX nucleic acid; (2) a host cell containing the vector; (3) an isolated
XX polypeptide or its fragment whose expression is inhibited by the
XX antisense nucleic acid; (4) an antibody capable of specifically binding
XX the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
XX proliferation or the activity of a gene in an operon required for
XX proliferation; (7) identifying a compound that influences the activity of
XX the gene product or that has an activity against a biological pathway
XX required for proliferation, or that inhibits cellular proliferation; (8)
XX identifying a gene required for cellular proliferation or the biological
XX pathway in which a proliferation-required gene or its gene product lies
XX or a gene on which the test compound that inhibits proliferation of an
XX organism acts; (9) manufacturing an antibiotic; (10) profiling a
XX compound's activity; (11) a culture comprising strains in which the gene
XX product is overexpressed or underexpressed; (12) determining the extent
XX to which each of the strains is present in a culture or collection of

CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 223 AA;
Query Match 89.7%; Score 26; DB 6; Length 223;
Best Local Similarity 71.4%; Pred. No. 1.5e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VMXVAEF 7
Db 93 VMNIAEF 99
|| :|||

RESULT 8
ID ABU36167 standard; protein; 224 AA.
XX AC ABU36167;
XX 19-JUN-2003 (first entry)
XX Protein encoded by Prokaryotic essential gene #21694.
XX Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX Mycoplasma pneumoniae.
XX WO200277183-A2.
XX 03-OCT-2002.
XX 21-MAR-2002; 2002WO-US009107.
XX 21-MAR-2001; 2001US-00815242.
XX 06-SEP-2001; 2001US-00948993.
XX 25-OCT-2001; 2001US-0342923P.
XX 08-FEB-2002; 2002US-00072851.
XX 06-MAR-2002; 2002US-0362699P.
XX (ELIT-) ELITRA PHARM INC.
XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
XX Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX WPI; 2003-029926/02.
XX N-PSDB; ACA40037.
XX New antisense nucleic acids, useful for identifying proteins or screening
XX for homologous nucleic acids required for cellular proliferation to
XX isolate candidate molecules for rational drug discovery programs.
XX Claim 25; SEQ ID NO 64091; 1766pp; English.
XX The invention relates to an isolated nucleic acid comprising any one of
XX the 6213 antisense sequences given in the specification where expression
XX of the nucleic acid inhibits proliferation of a cell. Also included are:
XX (1) a vector comprising a promoter operably linked to the nucleic acid
XX encoding a polypeptide whose expression is inhibited by the antisense
XX nucleic acid; (2) a host cell containing the vector; (3) an isolated
XX polypeptide or its fragment whose expression is inhibited by the
XX antisense nucleic acid; (4) an antibody capable of specifically binding
XX the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
XX proliferation or the activity of a gene in an operon required for

CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 XX Sequence 224 AA;

Query Match 89.7%; Score 26; DB 6; Length 224;
 Best Local Similarity 71.4%; Pred. No. 1.5e+02;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VMXVAEF 7
 Db 93 VMNIAEF 99

RESULT 9
 ADNA7550
 ID ADNA7550 standard; protein; 304 AA.

XX AC ADNA7550;

XX DT 01-JUL-2004 . (first entry)

DE Thermococcus kodakaraensis KOD1 protein sequence SeqID1428.

XX gene disruption; gene targeting; marker gene; transformation;
 KW homologous recombination; hyperthermostable archaeobacterium; KOD1;
 KW gene structure; gene function; enzyme activity; medicine;
 KW forensic science; food; drug inspection; molecular biology; immunology.

XX Thermococcus kodakaraensis.

XX WO2004022736-A1.

XX PD 18-MAR-2004.

XX PF 29-AUG-2003; 2003WO-IB003597.

XX PD 30-AUG-2002; 2002JP-00319011.

XX PA (NISC-) JAPAN SCI & TECHNOLOGY CORP.

XX PI Imanaka T, Atomi H;

XX WPI; 2004-257583/24.

XX Method for disrupting targeted gene in genome of organism particularly
 PT thermostable bacterium and with genome chips for analysis, applicable in
 PT studying gene structure and functions.

XX Claim 9; SEQ ID NO 1428; 598pp; Japanese.

XX This invention relates to a novel method for targeting disruption of an
 CC arbitrary gene in a genome of an organism which comprises providing the
 CC whole sequential data of the genome of such organism, selecting at least

CC 1 arbitrary region in the sequence, providing a vector that contains a
 CC sequence homologous with the selected region and a marker gene,
 CC transformation, and homologous recombination. The genome is preferably
 CC the genome of a hyperthermostable archaeobacterium, particularly
 CC *Thermococcus kodakaraensis* KOD1. The method is for targeting the
 CC disruption of a gene in the genome of an organism, which is applicable in
 CC studying gene structure and functions as well as enzyme activities of
 CC encoded proteins and useful in medicine, forensic science, food or drug
 CC inspection, molecular biology and immunology. With this method, the
 CC disruption of a gene at an arbitrary position in a genome can be achieved
 CC efficiently and reliably. The present sequence is that of a protein
 CC encoded by the genome of *Thermococcus kodakaraensis* which was derived
 CC using the method of the invention. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 XX Sequence 304 AA;

Query Match 89.7%; Score 26; DB 8; Length 304;
 Best Local Similarity 71.4%; Pred. No. 2.1e+02;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VMXVAEF 7
 Db 41 IMKVAEF 47

RESULT 10
 AAY34479
 ID AAY34479 standard; protein; 821 AA.

XX AC AAY34479;

XX DT 27-AUG-2003 (revised)

XX DT 20-MAR-2003 (revised)

XX DT 25-AUG-1999 (first entry)

DE Porphyromonas gingivalis protein PG21.

XX Porphyromonas gingivalis; PG; periodontal disease; gingivitis; vaccine;
 KW antigenic.

XX Porphyromonas gingivalis.

XX WO9929870-A1.

XX PD 17-JUN-1999.

XX PF 10-DEC-1998; 98WO-AU001023.

XX PR 10-DEC-1997; 97AU-00000839.

XX PR 31-DEC-1997; 97AU-00001182.

XX PR 30-JAN-1998; 98AU-00001546.

XX PR 10-MAR-1998; 98AU-00002264.

XX PR 09-APR-1998; 98AU-00002911.

XX PR 23-APR-1998; 98AU-00003128.

XX PR 05-MAY-1998; 98AU-00003338.

XX PR 22-MAY-1998; 98AU-00003654.

XX PR 29-JUL-1998; 98AU-00004917.

XX PR 30-JUL-1998; 98AU-00004963.

XX PR 04-AUG-1998; 98AU-00005028.

XX (CSLC-) CSL LTD.

XX Ross BC, Barr IG, Patterson MA, Agius CT, Rothel LJ, Margetts MB;
 PI Hocking DM, Webb EA;

XX WPI; 1999-385613/32.

XX N-PSDB; AAX31697.

XX Antigenic Porphyromonas gingivalis peptides for preventing gingivitis.

```

PS Claim 1; Page 460-461; 588pp; English.
XX
CC AAX91536 to AAX91801 encode two hundred and sixty six antigenic
CC Porphyromonas gingivalis (PG) polypeptide sequences given in AAY34318 to
CC AAY34583. AAX91802 to AAX91989 represent PCR primers used in the
CC isolation of the PG polypeptides. The PG polypeptides have antibacterial
CC activity with a vaccine mechanism of action. The PG polypeptides can be
CC used as vaccines especially against Porphyromonas gingivalis. Probes can
CC be used to detect Porphyromonas gingivalis in standard hybridisation
CC assays. Porphyromonas gingivalis is involved in periodontal disease
CC especially gingivitis. (Updated on 20-MAR-2003 to correct PR field.)
XX
SQ Sequence 821 AA;
Query Match 89.7%; Score 26; DB 2; Length 821;
Best Local Similarity 85.7%; Pred. No. 6.3e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
Db 565 VMPVAEF 571
||| |||

RESULT 11
AAY34354
ID AAY34354 standard; protein; 869 AA.
XX
AC AAY34354;
XX
DT 27-AUG-2003 (revised)
DT 20-MAR-2003 (revised)
DT 25-AUG-1999 (first entry)
XX
DE Porphyromonas gingivalis protein PG21.
XX
KW Porphyromonas gingivalis; PG; periodontal disease; gingivitis; vaccine;
KW antigenic.
XX
OS Porphyromonas gingivalis.
XX
PN W05929870-A1.
XX
PD 17-JUN-1999.
XX
PF 10-DEC-1998; 98WO-AU001023.
XX
PR 10-DEC-1997; 97AU-00000839.
PR 31-DEC-1997; 97AU-00001182.
PR 30-JAN-1998; 98AU-00001546.
PR 10-MAR-1998; 98AU-00002264.
PR 09-APR-1998; 98AU-00002911.
PR 23-APR-1998; 98AU-00003128.
PR 05-MAY-1998; 98AU-00003338.
PR 22-MAY-1998; 98AU-00003654.
PR 29-JUL-1998; 98AU-00004917.
PR 30-JUL-1998; 98AU-00004963.
PR 04-AUG-1998; 98AU-00005028.
XX
PA (CSLC-) CSL LTD.
XX
PI Rees BC, Barr IG, Patterson MA, Agius CT, Rothel LJ, Margetts MB;
PI Hocking DM, Webb EA;
XX
XX WPI; 1999-385613/32.
DR N-PSDB; AAX91572.
XX
PI Antigenic Porphyromonas gingivalis peptides for preventing gingivitis.
XX
Claim 1; Page 316-318; 588pp; English.
PS
XX AAX91536 to AAX91801 encode two hundred and sixty six antigenic
CC Porphyromonas gingivalis (PG) polypeptide sequences given in AAY34318 to
CC AAY34583. AAX91802 to AAX91989 represent PCR primers used in the
CC isolation of the PG polypeptides. The PG polypeptides have antibacterial
CC activity with a vaccine mechanism of action. The PG polypeptides can be
CC used as vaccines especially against Porphyromonas gingivalis. Probes can
CC be used to detect Porphyromonas gingivalis in standard hybridisation
CC assays. Porphyromonas gingivalis is involved in periodontal disease
CC especially gingivitis. (Updated on 20-MAR-2003 to correct PR field.)
XX
SQ Sequence 821 AA;
Query Match 89.7%; Score 26; DB 2; Length 821;
Best Local Similarity 85.7%; Pred. No. 6.3e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
Db 565 VMPVAEF 571
||| |||

RESULT 12
ADH85915
ID ADH85915 standard; protein; 194 AA.
XX
AC ADH85915;
XX
DT 22-APR-2004 (first entry)
XX
DE Enterococcus faecalis polypeptide #395.
XX
KW Enterococcus faecalis infection; transcription regulatory element;
KW antibacterial.
XX
OS Enterococcus faecalis.
XX
PN US6617156-B1.
XX
PD 09-SEP-2003.
XX
PF 13-AUG-1998; 98US-00134000.
XX
PR 15-AUG-1997; 97US-0055778P.
XX
PA (DOUC/) DOUCETTE-STAMM L A.
PA (BUSH/) BUSH D.
XX
PI Doucette-Stamm LA, Bush D;
XX
XX WPI; 2003-895394/82.
DR N-PSDB; ADH82510.
XX
XX New nucleic acid comprising a sequence encoding an Enterococcus faecalis
PT polypeptide, useful for preparing a composition for diagnosing or
PT treating E. faecalis infection.
XX
PS Disclosure; SEQ ID NO 3800; 193pp; English.
XX
XX The invention relates to Enterococcus faecalis polynucleotides and
CC polypeptides. The invention also relates to a recombinant expression
CC vector comprising a polynucleotide operably linked to a transcription
CC regulatory element, a cell comprising a recombinant vector, a method for
CC producing an E. faecalis polypeptide, an isolated nucleic acid comprising
CC a sequence not given in the specification, a recombinant vector
CC comprising the nucleic acid and a cell comprising the recombinant vector.
CC The polynucleotides can be used to detect the presence of E. faecalis in
CC a sample. The sequences are useful for preparing a composition for
CC diagnosing or treating Enterococcus faecalis infection. This sequence
CC represents an E. faecalis polypeptide of the invention.
XX
SQ Sequence 194 AA;
Query Match 86.2%; Score 25; DB 7; Length 194;
Best Local Similarity 71.4%; Pred. No. 2.4e+02;

```

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
|: |||||
Db 112 VLSVAEF 118

RESULT 13
ABU29329
ID ABU29329 standard; protein; 406 AA.
XX
AC ABU29329;
XX
DT 19-JUN-2003 (first entry)
DE
DE Protein encoded by Prokaryotic essential gene #14856.
XX
KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
OS Enterococcus faecalis.
XX
PN W0200277183-A2.
XX
PD 03-OCT-2002.
XX
PF 21-MAR-2002; 2002WO-US009107.
XX
PR 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 23-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JB, Carr GU, Yamamoto R, Forsyth RA, Xu HH;
XX
DR WPI; 2003-029926/02.
DR N-PSDB; ACA33199.
XX
XX New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
PS Claim 25; SEQ ID NO 57253; 1766pp; English.
XX
XX The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,

CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 406 AA;
Query Match 86.2%; Score 25; DB 6; Length 406;
Best Local Similarity 71.4%; Pred. No. 5.3e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
|: |||||
Db 111 VLSVAEF 117

RESULT 14
ADN48186
ID ADN48186 standard; protein; 583 AA.
XX
AC ADN48186;
XX
DT 01-JUL-2004 (first entry)
XX
DE Thermococcus kodakaraensis KOD1 protein sequence SeqID2064.
XX
KW gene disruption; gene targeting; marker gene; transformation;
KW homologous recombination; hyperthermostable archaeobacterium; KOD1;
KW gene structure; gene function; enzyme activity; medicine;
KW forensic science; food; drug inspection; molecular biology; immunology.
XX
OS Thermococcus kodakaraensis.
XX
PN W02004022736-A1.
XX
PD 18-MAR-2004.
XX
PF 29-AUG-2003; 2003WO-IB003597.
XX
PR 30-AUG-2002; 2002JP-00319011.
XX
PA (NISC-) JAPAN SCI & TECHNOLOGY CORP.
XX
PI Imanaka T, Atomi H;
XX
DR WPI; 2004-257583/24.
XX
XX Method for disrupting targeted gene in genome of organism particularly
PT thermostable bacterium and with genome chips for analysis, applicable in
PT studying gene structure and functions.
XX
PS Claim 9; SEQ ID NO 2064; 598pp; Japanese.
XX
XX This invention relates to a novel method for targeting disruption of an
CC arbitrary gene in a genome of an organism which comprises providing the
CC whole sequential data of the genome of such organism, selecting at least
CC 1 arbitrary region in the sequence, providing a vector that contains a
CC sequence homologous with the selected region and a marker gene,
CC transformation, and homologous recombination. The genome is preferably
CC the genome of a hyperthermostable archaeobacterium, particularly
CC Thermococcus kodakaraensis KOD1. The method is for targeting the
CC disruption of a gene in the genome of an organism, which is applicable in
CC studying gene structure and functions as well as enzyme activities of
CC encoded proteins and useful in medicine, forensic science, food or drug
CC inspection, molecular biology and immunology. With this method, the
CC disruption of a gene at an arbitrary position in a genome can be achieved
CC efficiently and reliably. The present sequence is that of a protein
CC encoded by the genome of Thermococcus kodakaraensis which was derived
CC using the method of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences


```

QY      1 VMXVAEF 7
DB      53 VMDVAEY 59

RESULT 17
AAU56778
ID AAU56778 standard; protein; 75 AA.
XX
AC AAU56778;
XX
DT 27-FEB-2002 (first entry)
XX
DE Propionibacterium acnes immunogenic protein #17674.
XX
KW SAPHO syndrome; synovitis; acne; pustulosis; hypertension; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant.
XX
OS Propionibacterium acnes.
XX
PN WO200181581-A2.
XX
PD 01-NOV-2001.
XX
PF 20-APR-2001; 2001WO-US012865.
XX
PR 21-APR-2000; 2000US-0199047P.
PR 02-JUN-2000; 2000US-0208841P.
PR 07-JUL-2000; 2000US-0216747P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
PI L'Maisonneuve J, Zhang Y, Jen S, Carter D;
XX
DR WPI; 2001-616774/71.
DR N-PSDB; AAS59578.
XX
PT Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris.
XX
PS Example 1; SEQ ID NO 17973; 1069pp; English.
XX
CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hypertension and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 75 AA;

Query Match      82.8%; Score 24; DB 4; Length 75;
Best Local Similarity 71.4%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 VMXVAEF 7
DB      1 VMAVCEF 7

RESULT 18
ABM53297
ID ABM53297 standard; protein; 75 AA.
XX
AC ABM53297;
XX
DT 20-OCT-2003 (first entry)
XX
DE Propionibacterium acnes predicted ORF-encoded polypeptide #17973.
XX
KW Acne vulgaris; antiseborrheic; dermatological; antibacterial;
KW immunostimulant; immune response; vaccine.
XX
OS Propionibacterium acnes.
XX
PN WO2003033515-A1.
XX
PD 24-APR-2003.
XX
PF 11-OCT-2002; 2002WO-US032727.
XX
PR 15-OCT-2001; 2001US-00978825.
XX
PA (CORI-) CORIXA CORP.
XX
PI Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;
PI Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;
PI Barth B, Vallieve-Douglas J;
XX
DR WPI; 2003-381789/36.
DR N-PSDB; ACF64507.
XX
PT New Propionibacterium acnes polypeptides and polynucleotides encoding the
PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
PT or for stimulating an immune response specific for a P. acnes protein.
XX
PS Example 1; SEQ ID NO 17973; 1481pp; English.
XX
CC The invention relates to an isolated polynucleotide (ACF64435-ACF64733)
CC encoding a Propionibacterium acnes protein. The invention also relates to
CC polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to
CC immunogenic fragments of P. acnes polypeptides. The invention
CC additionally encompasses expression vectors and host cells comprising a
CC polynucleotide of the invention; antibodies against polypeptides of the
CC invention; fusion proteins comprising a polypeptide of the invention; a
CC method for stimulating an immune response specific for a P. acnes
CC polypeptide and an isolated T cell population comprising T cells prepared
CC via this method; a vaccine composition (comprising P. acnes polypeptides,
CC polynucleotides, antibodies, fusion proteins, T cell populations, or
CC antigen-presenting cells that express the polypeptide); a method and kit
CC for detecting or determining the presence or absence of P. acnes in a
CC patient; and a method for inhibiting the development of P. acnes in a
CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion
CC proteins, T cell populations or antigen-presenting cells that express the
CC polypeptides are useful for diagnosing, preventing or treating acne
CC vulgaris, or for stimulating an immune response specific for a P. acnes
CC protein. The polynucleotides can also be used as probes or primers for
CC nucleic acid hybridisation. The vaccine composition is useful for the
CC stimulation of an immune response against P. acnes, or for treating acne,
CC and the kit is useful for performing a diagnostic assay. The present
CC sequence represents a polypeptide predicted to be encoded by an ORF (open
CC reading frame) contained within the P. acnes polynucleotides of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 75 AA;

```

Query Match 82.8%; Score 24; DB 6; Length 75;
 Best Local Similarity 71.4%; Pred. No. 1.6e+02;
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
 || || ||
 DB 1 VMAVCEF 7

RESULT 19
 ABP75586
 ID ABP75586 standard; protein; 97 AA.
 XX
 AC ABP75586;
 XX
 DT 10-FEB-2003 (first entry)
 XX
 DE Human secretory polypeptide SPTM SEQ ID NO 770.
 XX
 KW Human; SPTM; autoimmune disorder; inflammatory disorder; AIDS; anaemia;
 KW asthma; Crohn's disease; neurological disorder; epilepsy; cancer;
 KW Huntington's disease; Alzheimer's disease; Creutzfeldt-Jakob disease;
 KW multiple sclerosis; Parkinson's disease; cell proliferative disorder;
 KW anti-inflammatory; immunosuppressive; neuroprotective; neurotropic;
 KW neuroleptic; anticonvulsant; cytostatic; antiparkinsonian; anxiolytic;
 KW antipsoriatic; antianaemic; anti-HIV; human immunodeficiency virus;
 KW secretory polynucleotide; secretory protein.
 XX
 OS Homo sapiens.
 XX
 XX
 PN WO200283876-A2.
 XX
 PD 24-OCT-2002.
 XX
 XX
 PF 27-MAR-2002; 2002WO-US009921.
 XX
 XX
 PR 29-MAR-2001; 2001US-0280067P.
 PR 29-MAR-2001; 2001US-0280068P.
 PR 16-MAY-2001; 2001US-0291280P.
 PR 17-MAY-2001; 2001US-0291829P.
 PR 17-MAY-2001; 2001US-0291849P.
 PR 19-JUN-2001; 2001US-0299428P.
 PR 20-JUN-2001; 2001US-029976P.
 PR 20-JUN-2001; 2001US-0300001P.
 XX
 XX (INCYTE GENOMICS INC.
 XX
 XX Daffo A, Jones AL, Tran AB, Dahl CR, Gietzen D, Chinn J;
 PI Dufour GE, Hillman JL, Yu JY, Tason O, Yap PE, Ameshey SR;
 PI Daugherty SC, Dam TC, Liu TF, Nguyen DA, Kleefeld Y, Gerstin EH;
 PI Peralta CH, David MH, Lewis SA, Chen AJ, Panzer SR, Harris B;
 PI Flores V, Marwaha R, Lo A, Lan RV, Urashka MF;
 XX WPI; 2003-075543/07.
 DR N-PSDB; ABZ36031.
 XX
 XX New human secretory proteins and polynucleotides, useful for diagnosing,
 PT treating or preventing autoimmune/inflammatory disorders (e.g. AIDS),
 PT neurological disorders (e.g. Alzheimer's), or cell proliferations or
 PT cancers.
 XX
 XX Claim 27; SEQ ID NO 770; 458pp + Sequence Listing; English.
 XX
 XX The invention relates to a secretory polynucleotide (designated sptm)
 CC comprising any of 567 polynucleotide sequences (ABZ35837-ABZ36403), a
 CC naturally occurring polynucleotide sequence at least 90 % identical to
 CC the polynucleotide sequence, a polynucleotide complementary to them or an
 CC RNA equivalent of them. The polypeptide or polynucleotide are useful for
 CC treating, preventing or diagnosing a disease or condition associated with
 CC the expression of functional SPTM. These are particularly useful for
 CC diagnosing, treating or preventing autoimmune/inflammatory disorders
 CC (e.g. acquired immunodeficiency syndrome, anaemia, asthma or Crohn's

CC disease), neurological disorders (e.g. epilepsy, Huntington's disease,
 CC dementia, stroke, Alzheimer's disease, Creutzfeldt-Jakob disease,
 CC multiple sclerosis, cerebral palsy, Parkinson's disease, anxiety,
 CC schizophrenia or amnesia), or cell proliferative disorders (e.g.
 CC psoriasis, polycythemia vera, or cancers including adenocarcinoma,
 CC leukaemia, lymphoma, melanoma, myeloma, sarcoma or cancers of the brain,
 CC breast, cervix or prostate). The present sequence is one of the SPTM
 CC proteins of the invention (ABP75384-ABP75962). Note: The sequence data
 CC for this patent did not form part of the printed specification, but was
 CC obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 97 AA;
 Query Match 82.8%; Score 24; DB 6; Length 97;
 Best Local Similarity 83.3%; Pred. No. 2.1e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 MXVAEF 7
 | ||||
 DB 1 MSVAEF 6

RESULT 20
 AAM41985
 ID AAM41985 standard; protein; 132 AA.
 XX
 AC AAM41985;
 XX
 DT 22-OCT-2001 (first entry)
 XX
 DE Human polypeptide SEQ ID NO 6916.
 XX
 KW Human; neurotropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 KW leukaemia.
 XX
 XX Homo sapiens.
 OS
 XX
 PN WO200153312-A1.
 XX
 XX 26-JUL-2001.
 PD
 XX
 PF 26-DEC-2000; 2000WO-US034263.
 XX
 XX 23-DEC-1999; 99US-00471275.
 PR 21-JAN-2000; 2000US-00488725.
 PR 25-APR-2000; 2000US-0052317.
 PR 20-JUN-2000; 2000US-00598042.
 PR 19-JUL-2000; 2000US-00620312.
 PR 03-AUG-2000; 2000US-00653450.
 PR 14-SEP-2000; 2000US-00662191.
 PR 19-OCT-2000; 2000US-00693036.
 PR 29-NOV-2000; 2000US-00727344.
 XX
 XX (HYSEQ-) HYSEQ INC.
 XX
 XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
 PI Zhou P, Goodrich R, Drmanac RT;
 XX WPI; 2001-442253/47.
 DR N-PSDB; AAI61141.
 XX
 XX Novel nucleic acids and polypeptides, useful for treating disorders such
 PT as central nervous system injuries.
 XX
 XX Example 2; SEQ ID NO 6916; 10078pp; English.
 PS
 XX The invention relates to human nucleic acids (AAI57798-AAI61369) and the

CC encoded polypeptides (AAM38642-AAM42213) with nootropic,
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: Immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening
 CC assays for receptor activity, arthritis and inflammation, leukaemias and
 CC C.N.S disorders. Note: The sequence data for this patent did not form
 CC part of the printed specification
 XX
 SQ Sequence 132 AA;
 Query Match 82.8%; Score 24; DB 4; Length 132;
 Best Local Similarity 71.4%; Pred. No. 2.9e+02;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 VMXVAEF 7
 : |||||
 Db 64 LMVVAEF 70
 RESULT 21
 AAM41986
 ID AAM41986 standard; protein; 132 AA.
 AC AAM41986;
 XX
 DT 22-OCT-2001 (first entry)
 XX
 DE Human polypeptide SEQ ID NO 6917.
 XX
 KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 KW leukaemia.
 XX
 OS Homo sapiens.
 XX
 PN WO200153312-A1.
 XX
 PD 26-JUL-2001.
 XX
 PF 26-DEC-2000; 2000WO-US034263.
 XX
 PR 21-DEC-1999; 99US-00471275.
 PR 21-JAN-2000; 2000US-00488725.
 PR 25-APR-2000; 2000US-00552317.
 PR 20-JUN-2000; 2000US-00598042.
 PR 19-JUL-2000; 2000US-00620312.
 PR 03-AUG-2000; 2000US-00653450.
 PR 14-SEP-2000; 2000US-00662191.
 PR 19-OCT-2000; 2000US-00693036.
 PR 29-NOV-2000; 2000US-00727344.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
 PI Zhou P, Goodrich R, Drmanac RT;
 XX
 DR WPI; 2001-442253/47.
 DR N-PSDB; AAI61142.
 XX
 PT Novel nucleic acids and polypeptides, useful for treating disorders such
 PT as central nervous system injuries.
 XX

PS Example 2; SEQ ID NO 6917; 10078pp; English.
 XX
 CC The invention relates to human nucleic acids (AAI57798-AAI61369) and the
 CC encoded polypeptides (AAM38642-AAM42213) with nootropic,
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: Immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemias and
 CC C.N.S disorders. Note: The sequence data for this patent did not form
 CC part of the printed specification
 XX
 SQ Sequence 132 AA;
 Query Match 82.8%; Score 24; DB 4; Length 132;
 Best Local Similarity 71.4%; Pred. No. 2.9e+02;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 VMXVAEF 7
 : |||||
 Db 64 LMVVAEF 70
 RESULT 22
 AAM41984
 ID AAM41984 standard; protein; 132 AA.
 AC AAM41984;
 XX
 DT 22-OCT-2001 (first entry)
 XX
 DE Human polypeptide SEQ ID NO 6915.
 XX
 KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 KW leukaemia.
 XX
 OS Homo sapiens.
 XX
 PN WO200153312-A1.
 XX
 PD 26-JUL-2001.
 XX
 PF 26-DEC-2000; 2000WO-US034263.
 XX
 PR 23-DEC-1999; 99US-00471275.
 PR 21-JAN-2000; 2000US-00488725.
 PR 25-APR-2000; 2000US-00552317.
 PR 20-JUN-2000; 2000US-00598042.
 PR 19-JUL-2000; 2000US-00620312.
 PR 03-AUG-2000; 2000US-00653450.
 PR 14-SEP-2000; 2000US-00662191.
 PR 19-OCT-2000; 2000US-00693036.
 PR 29-NOV-2000; 2000US-00727344.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
 PI Zhou P, Goodrich R, Drmanac RT;
 XX
 DR WPI; 2001-442253/47.
 DR N-PSDB; AAI61140.
 XX

PT Novel nucleic acids and polypeptides, useful for treating disorders such
PT as central nervous system injuries.

PS Example 2; SEQ ID NO 6915; 10078pp; English.

XX
XX The invention relates to human nucleic acids (AAI57798-AAI61369) and the
CC encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC immunosuppressant and cytoskeletal activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders. Note: The sequence data for this patent did not form
CC part of the printed specification

XX SQ Sequence 132 AA;

Query Match 82.8%; Score 24; DB 4; Length 132;
Best Local Similarity 71.4%; Pred. No. 2.9e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
:| |||||
Db 64 LMVVAEF 70

RESULT 23

ID ABG98409
ID ABG98409 standard; protein; 145 AA.

AC ABG98409;

XX 16-JAN-2003 (first entry)

XX Haloarcula marsimortui 50S ribosomal protein L13.

XX Ribosome; three dimensional structure; Thermus thermophilus;
KW x-ray crystallography; protein coordinate data.

XX Haloarcula marsimortui.

XX WO200246392-A2.

XX 13-JUN-2002.

XX 10-DEC-2001; 2001WO-US047975.

XX 09-DEC-2000; 2000US-0254603P.

PR 22-MAR-2001; 2001US-0278013P.

PR 30-MAY-2001; 2001US-0294394P.

XX (REGC) UNIV CALIFORNIA.

XX Noller HF, Cate JHD, Yusupov MM, Yusupova GZ, Baucom AE;
PI Lancaster L, Dallas A, Lieberman K;

XX WPI; 2002-713237/77.

XX Novel three dimensional structure of Thermus thermophilus 70S ribosome
PT resolved using x-ray crystallography upto 5.5 Angstroms resolution,
PT useful for screening and designing compounds that alter ribosome
PT function.

XX Disclosure; Page 502; 525pp; English.

XX This invention relates to a three dimensional structure of Thermus
CC thermophilus 70S ribosome resolved using x-ray crystallography up to 5.5

CC Angstrom resolution. The invention also discloses a method for
CC identifying a compound that binds to a bacterial 70S ribosome or its
CC portion. The method comprises designing a compound based upon a three-
CC dimensional structure of the bacterial 70S ribosome, where the structure
CC co-ordinates are given in the specification, contacting the compound with
CC the bacterial 70S ribosome or its portion and determining whether the
CC compound binds to the bacterial 70S ribosome or its portion. The 70S
CC ribosome 3D structure of the invention can be used to screen and design
CC compounds that bind to sites on the 70S ribosome and that alter the
CC ribosome function. The present sequence represents a ribosomal protein
CC sequence used in the method of the invention

XX SQ Sequence 145 AA;

Query Match 82.8%; Score 24; DB 5; Length 145;
Best Local Similarity 83.3%; Pred. No. 3.2e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 MXVAEF 7
| |||||
Db 1 MSVAEF 6

RESULT 24

ABU24614
ID ABU24614 standard; protein; 191 AA.

XX AC ABU24614;

XX 19-JUN-2003 (first entry)

XX Protein encoded by Prokaryotic essential gene #10141.

XX Antisense; prokaryotic essential gene; cell proliferation; drug design.

OS Clostridium botulinum.

XX WO200277183-A2.

XX 03-OCT-2002.

XX 21-MAR-2002; 2002WO-US009107.

XX 21-MAR-2001; 2001US-00815242.

PR 06-SEP-2001; 2001US-00948993.

PR 25-OCT-2001; 2001US-0342923P.

PR 08-FEB-2002; 2002US-00072851.

PR 06-MAR-2002; 2002US-0362699P.

XX (ELIT-) ELITRA PHARM INC.

XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX WPI; 2003-029926/02.

XX N-PSDB; ACA28484.

XX New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.

XX Claim 25; SEQ ID NO 52538; 1766pp; English.

XX The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for

CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 191 AA;

Query Match 82.8%; Score 24; DB 6; Length 191;
 Best Local Similarity 57.1%; Pred. No. 4.3e+02;
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
 Db 131 LMAIAEF 137

RESULT 25

ABU23979
 ID ABU23979 standard; protein; 195 AA.

XX AC ABU23979;

XX 19-JUN-2003 (first entry)

DE Protein encoded by Prokaryotic essential gene #9506.

XX Antisense; prokaryotic essential gene; cell proliferation; drug design.

XX OS Clostridium acetobutylicum.

XX PN W0200277183-A2.

XX PD 03-OCT-2002.

XX PF 21-MAR-2002; 2002WO-US009107.

XX PR 21-MAR-2001; 2001US-00815242.

PR 06-SEP-2001; 2001US-00948993.

PR 25-OCT-2001; 2001US-0342923P.

PR 08-FEB-2002; 2002US-00072851.

PR 06-MAR-2002; 2002US-0362699P.

XX PA (ELIT-) ELITRA PHARM INC.

XX PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;

PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX DR WPI; 2003-029926/02.

DR N-PSDB; ACA27849.

XX New antisense nucleic acids, useful for identifying proteins or screening

PT for homologous nucleic acids required for cellular proliferation to

PT isolate candidate molecules for rational drug discovery programs.

XX PS Claim 25; SEQ ID NO 51903; 1766pp; English.

XX

CC The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 195 AA;

Query Match 82.8%; Score 24; DB 6; Length 195;
 Best Local Similarity 57.1%; Pred. No. 4.4e+02;
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
 Db 131 LMAIAEF 137

RESULT 26

ADJ48536

ID ADJ48536 standard; protein; 195 AA.

XX AC ADJ48536;

XX 06-MAY-2004 (first entry)

DE Oil-associated gene related protein #36.

XX oil-associated gene; transgenic; enhanced seed oil; vegetable oil.

XX OS Unidentified.

XX PN US2004025202-A1.

XX PD 05-FEB-2004.

XX PF 14-MAR-2003; 2003US-00389566.

XX PR 15-MAR-2002; 2002US-0365301P.

PR 26-JUN-2002; 2002US-0391786P.

PR 26-JUN-2002; 2002US-0392018P.

XX (LAUR/) LAURIE C C.

PA (RAVA/) RAVANELLO M.

PA (SAVA/) SAUSAGE T.

PA (LEDE/) LEDEUX J R.

PA (ROGE/) ROGERS J A.

XX

PI Laurie CC, Ravanello M, Savage T, Ledeaux JR, Rogers JA;
 XX WPI; 2004-142683/14.
 XX
 XX Novel recombinant DNA construct comprising a promoter functional in
 PT plants operably linked to an oil-associated gene for producing transgenic
 PT plant seed.
 XX
 PS Example 3; SEQ ID NO 540; 22pp; English.
 XX
 CC The invention relates to a recombinant DNA construct comprising a
 CC promoter functional in plants operably linked to an oil-associated gene.
 CC The construct is useful for transgenic plant seed which has in its genome
 CC the construct, that is functional in the plant to transcribe the oil-
 CC associated gene. The transgenic plant seed grows into a plant having
 CC enhanced seed oil as compared to wild type. The construct is useful for
 CC producing hybrid maize seed. The transgenic plant seed is useful for
 CC producing vegetable oil. The present sequence represents the amino acid
 CC sequence of an oil-associated gene related protein.
 XX
 SQ Sequence 195 AA;
 Query Match 82.8%; Score 24; DB 8; Length 195;
 Best Local Similarity 71.4%; Pred. No. 4.4e+02;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 VMKVAEF 7
 || |||
 Db 73 VMLVAEY 79
 XX
 RESULT 27
 AAY85820
 ID AAY85820 standard; protein; 225 AA.
 XX
 AC AAY85820;
 XX
 DT 10-APR-2000 (first entry)
 XX
 DE S. pneumoniae derived protein #29.
 XX
 KW Treatment; prevention; disease; diagnosis; gene therapy; screening;
 KW bacterial; antimicrobial; antibiotic; pathogenesis; infection.
 XX
 OS Streptococcus pneumoniae.
 XX
 PN WO9806734-A1.
 XX
 PD 19-FEB-1998.
 XX
 PF 15-AUG-1997; 97WO-US014436.
 XX
 PR 16-AUG-1996; 96US-0024022P.
 XX
 XX (SMIK) SMITHKLINE BEECHAM CORP.
 XX
 PI Black MT, Hodgson JE, Knowles DJC, Lonetto MA, Nicholas RO;
 PI Stodola RK;
 XX
 DR WPI; 1998-159452/14.
 DR N-PSDB; AAZ96199.
 XX
 XX Streptococcus pneumoniae proteins and related DNA - useful for screening
 PT compounds for antibacterial activity.
 PS
 PS Claim 5; Page 338-339; 640pp; English.
 XX
 CC This invention describes novel isolated Streptococcus pneumoniae
 CC polynucleotides (see AAZ96173-236494) and their encoded proteins (see
 CC AAY85792-Y86182). The DNA, vectors and host cells described in the method
 CC of the invention are useful for the recombinant expression of the
 CC polypeptides. The polypeptides are useful for treatment or prevention of
 CC disease, or diagnosis of disease related to expression or activity of

CC such a polypeptide. They can also be used to screen for compounds which
 CC interact with and inhibit or activate such a polypeptide. The
 CC polypeptides (or DNA encoding them, via gene therapy) are also useful for
 CC inducing an immunological response in a mammal. The antagonists are
 CC useful to inhibit such bacterial polypeptides. The polypeptides are
 CC particularly useful to identify antimicrobial compounds and antibiotics.
 CC They are also useful to determine their role in pathogenesis of
 CC infection, dysfunction and disease
 XX
 SQ Sequence 225 AA;
 Query Match 82.8%; Score 24; DB 2; Length 225;
 Best Local Similarity 71.4%; Pred. No. 5.1e+02;
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 VMKVAEF 7
 || |||
 Db 115 VMSAAEF 121
 XX
 RESULT 28
 AAU50802
 ID AAU50802 standard; protein; 236 AA.
 XX
 AC AAU50802;
 XX
 DT 27-FEB-2002 (first entry)
 XX
 DE Propionibacterium acnes immunogenic protein #11698.
 XX
 KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
 KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 KW dermatological; osteopathic; neuroprotectant.
 XX
 OS Propionibacterium acnes.
 XX
 PN WO200181581-A2.
 XX
 PD 01-NOV-2001.
 XX
 XX 20-APR-2001; 2001WO-US012865.
 PF
 XX 21-APR-2000; 2000US-0199047P.
 PR
 PR 02-JUN-2000; 2000US-0208941P.
 PR 07-JUL-2000; 2000US-0216747P.
 XX
 XX (CORI-) CORIXA CORP.
 PA
 XX Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
 PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
 XX
 DR WPI; 2001-616774/71.
 DR N-PSDB; AAS95549.
 XX
 XX Propionibacterium acnes polypeptides and nucleic acids useful for
 PT vaccinating against and diagnosing infections, especially useful for
 PT treating acne vulgaris.
 XX
 PS Example 1; SEQ ID NO 11997; 1069pp; English.
 XX
 CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
 CC polypeptides. The proteins and their associated DNA sequences are used in
 CC the treatment, prevention and diagnosis of medical conditions caused by
 CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
 CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
 CC P. acnes is also involved in infections of bone, joints and the central
 CC nervous system, however it is particularly involved in the inflammatory
 CC lesions associated with acne vulgaris. A method for detecting the
 CC presence or absence of P. acnes in a patient comprises contacting a
 CC sample with a binding agent that binds to the proteins of the invention
 CC and determining the amount of bound protein in the sample. The
 CC polypeptides may be used as antigens in the production of antibodies

CC specific for P. acnes proteins. These antibodies can be used to
 CC downregulate expression and activity of P. acnes polypeptides and
 CC therefore treat P. acnes infections. The antibodies may also be used as
 CC diagnostic agents for determining P. acnes presence, for example, by
 CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
 CC this patent did not form part of the printed specification, but was
 CC obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 236 AA;

Query Match 82.8%; Score 24; DB 4; Length 236;
 Best Local Similarity 71.4%; Pred. No. 5.4e+02;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 VMXVAEF 7
 |||:||||
 Db 169 VMELAEF 175

RESULT 29
 ABM47321
 ID ABM47321 standard; protein; 236 AA.
 XX AC ABM47321;
 XX AC
 XX 20-OCT-2003 (first entry)
 XX DE Propionibacterium acnes predicted ORF-encoded polypeptide #11997.
 XX DE
 XX DE Acne vulgaris; antiseborrheic; dermatological; antibacterial;
 XX KW immunostimulant; immune response; vaccine.
 XX OS Propionibacterium acnes.
 XX OS
 XX PN WO2003033515-A1.
 XX PN
 XX PD 24-APR-2003.
 XX PD
 XX PF 11-OCT-2002; 2002WO-US032727.
 XX PF
 XX PR 15-OCT-2001; 2001US-00978825.
 XX PR
 XX PA (CORI-) CORIXA CORP.
 XX PA
 XX PI Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;
 XX PI Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;
 XX PI Barth B, Vallie-Douglas J;
 XX DR WPI; 2003-381789/36.
 XX DR N-PSDB; ACF64478.
 XX DR
 XX PT New Propionibacterium acnes polypeptides and polynucleotides encoding the
 XX PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
 XX PT or for stimulating an immune response specific for a P. acnes protein.
 XX PS Example 1; SEQ ID NO 11997; 1481pp; English.
 XX CC The invention relates to an isolated polynucleotide (ACF64435-ACF64733)
 CC encoding a Propionibacterium acnes protein. The invention also relates to
 CC polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to
 CC immunogenic fragments of P. acnes polypeptides. The invention
 CC additionally encompasses expression vectors and host cells comprising a
 CC polynucleotide of the invention; antibodies against polypeptides of the
 CC invention; fusion proteins comprising a polypeptide of the invention; a
 CC method for stimulating an immune response specific for a P. acnes
 CC polypeptide and an isolated T cell population comprising T cells prepared
 CC via this method; a vaccine composition (comprising P. acnes polypeptides,
 CC polynucleotides, antibodies, fusion proteins, T cell populations, or
 CC antigen-presenting cells that express the polypeptide); a method and kit
 CC for detecting or determining the presence or absence of P. acnes in a
 CC patient; and a method for inhibiting the development of P. acnes in a
 CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion

CC proteins, T cell populations or antigen-presenting cells that express the
 CC polypeptides are useful for diagnosing, preventing or treating acne
 CC vulgaris, or for stimulating an immune response specific for a P. acnes
 CC protein. The polynucleotides can also be used as probes or primers for
 CC nucleic acid hybridisation. The vaccine composition is useful for the
 CC stimulation of an immune response against P. acnes, or for treating acne,
 CC and the kit is useful for performing a diagnostic assay. The present
 CC sequence represents a polypeptide predicted to be encoded by an ORF (open
 CC reading frame) contained within the P. acnes polynucleotides of the
 CC invention. Note: The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 236 AA;

Query Match 82.8%; Score 24; DB 6; Length 236;
 Best Local Similarity 71.4%; Pred. No. 5.4e+02;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
 |||:||||
 Db 169 VMELAEF 175
 RESULT 30
 ABU26266
 ID ABU26266 standard; protein; 238 AA.
 XX AC ABU26266;
 XX AC
 XX 19-JUN-2003 (first entry)
 XX DT
 XX DE Protein encoded by Prokaryotic essential gene #11793.
 XX DE
 XX KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
 XX OS Corynebacterium diphtheriae.
 XX OS
 XX PN WO200277183-A2.
 XX PN
 XX PD 03-OCT-2002.
 XX PD
 XX PF 21-MAR-2002; 2002WO-US009107.
 XX PF
 XX PR 21-MAR-2001; 2001US-00815242.
 XX PR 06-SEP-2001; 2001US-00948993.
 XX PR 25-OCT-2001; 2001US-0342923P.
 XX PR 08-FEB-2002; 2002US-00072851.
 XX PR 06-MAR-2002; 2002US-0362699P.
 XX PR
 XX PA (ELIT-) ELITRA PHARM INC.
 XX PA
 XX PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
 XX PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX DR WPI; 2003-029926/02.
 XX DR N-PSDB; ACA30136.
 XX DR
 XX PT New antisense nucleic acids, useful for identifying proteins or screening
 XX PT for homologous nucleic acids required for cellular proliferation to
 XX PT isolate candidate molecules for rational drug discovery programs.
 XX PS Claim 25; SEQ ID NO 54190; 1766pp; English.
 XX CC The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular

CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 CC
 XX
 SQ Sequence 238 AA;

Query Match 82.8%; Score 24; DB 6; Length 238;
 Best Local Similarity 57.1%; Pred. No. 5.4e+02;
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
 | : : : : :
 Db 72 VLSTAEF 78

RESULT 31
 AAY36930
 ID AAY36930 standard; protein; 254 AA.

AC AAY36930;

XX 07-OCT-1999 (first entry)

XX Amino acid sequence of a Chlamydia trachomatis protein.

DE Vaccine; eye disease; conventional trachoma; nonendemic trachoma;
 KW paratrachoma; inclusion conjunctivitis; genital disease; perinephatitis;
 KW nongonococcal urethritis; epididymitis; cervicitis; salpingitis;
 KW bartholinitis; pneumopathy; venereal lymphogranulomatosis.

XX Chlamydia trachomatis.

OS Chlamydia trachomatis.

XX WO9928475-A2.

XX 10-JUN-1999.

XX 27-NOV-1998; 98WO-IB001939.

XX 28-NOV-1997; 97FR-00015041.

PR 17-DEC-1997; 97FR-00016034.

PR 04-NOV-1998; 98US-0107077P.

XX (GEST) GENSET.

XX Griffais R;

XX WPI; 1999-371125/31.

XX Genome sequence of Chlamydia trachomatis.

PS Disclosure; Page 786; 1755pp; English.

XX AAY36754-Y37949 are encoded by open reading frames (ORFs) of the genome
 CC of Chlamydia trachomatis (see AA201425). The polypeptides can be used as

CC vaccines against Chlamydia trachomatis. Antisense and ribozyme sequences
 CC can also be used to control growth of the microorganism. Chlamydia
 CC trachomatis is responsible for a large number of diseases, e.g. eye
 CC diseases such as conventional trachoma, nonendemic trachoma,
 CC paratrachoma, and inclusion conjunctivitis; genital diseases such as
 CC nongonococcal urethritis, epididymitis, cervicitis, salpingitis,
 CC perinephatitis, bartholinitis; pneumopathy in breast feeding infants; and
 CC venereal lymphogranulomatosis. The polypeptides of the invention may be
 CC of use in treating these diseases

XX Sequence 254 AA;

Query Match 82.8%; Score 24; DB 2; Length 254;

Best Local Similarity 71.4%; Pred. No. 5.8e+02;

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
 | : : : : :
 Db 6 VMIVADP 12

RESULT 32

ABG18334

ID ABG18334 standard; protein; 286 AA.

XX AC ABG18334;

XX 18-FEB-2002 (first entry)

XX Novel human diagnostic protein #18325.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.

XX Homo sapiens.

XX WO200175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US008631.

XX 31-MAR-2000; 2000US-00540217.

PR 23-AUG-2000; 2000US-00649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

XX N-PSDB; AAS82521.

XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.

XX Claim 20; SEQ ID NO 48693; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations

CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
 CC amino acid sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 286 AA;

Query Match 82.8%; Score 24; DB 4; Length 286;
 Best Local Similarity 71.4%; Pred. No. 6.6e+02;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 VMXVAEF 7
 || || || ||
 Db 23 VMQVQAF 29

RESULT 33
 ABB81185
 ID ABB81185 standard; protein; 332 AA.
 XX
 AC ABB81185;
 XX
 DT 25-NOV-2002 (first entry)
 XX
 DE VAB-1 protein sequence.
 XX
 KW Receptor tyrosine kinase; RTK; kinase domain; cytostatic; antiarthritic;
 KW antiinflammatory; immunosuppressive; antirheumatic; virucide; nootropic;
 KW neuroprotective; cerebroprotective; antiparkinsonian; dermatological;
 KW nephrotropic; tranquilizer; vulnerary; anticonvulsant; VAB-1.
 XX
 OS Unidentified.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 1..332
 FT /note= "residues Xaa are not clearly indicated in the
 FT specification"
 XX
 PN WO200261055-A2.
 XX
 PD 08-AUG-2002.
 XX
 PF 31-JAN-2002; 2002WO-CA000114.
 XX
 PR 31-JAN-2001; 2001US-0265510P.
 XX
 PA (MOUN) MOUNT SINAI HOSPITAL.
 XX
 PI Sicheiri F, Wybenga-Groot L, Pawson T;
 XX
 DR WPI; 2002-643365/69.
 XX
 PT Novel isolated binding pocket of receptor tyrosine kinase that regulates
 PT the kinase domain of the receptor, useful for identifying modulator of
 PT the receptor for treating lymphoproliferative conditions.
 XX
 PS Disclosure; Fig 1; 116pp; English.
 XX
 CC The invention relates to an isolated binding pocket (I) of a receptor
 CC tyrosine kinase (RTK) that regulates the kinase domain of RTK. A crystal
 CC (II) comprising a binding pocket of an RTK that regulates the kinase
 CC domain of the RTK, or comprising a juxtamembrane region and/or kinase
 CC domain of an RTK or its part, or formed by a juxtamembrane region and a
 CC kinase region of an RTK in an autoinhibited state and a model (III) of
 CC (I) made using (I); are useful for determining the secondary and/or
 CC tertiary structure of a polypeptide, or for screening for a ligand
 CC capable of binding to a binding pocket and/or inhibiting or enhancing the
 CC atomic contacts of interactions in a binding pocket. (I) is useful for
 CC identifying a modulator of an RTK. (II) is useful for designing,
 CC modelling, identifying, evaluating and/or synthesizing mimetics of a

CC binding pocket, or ligands that associate with the binding pocket, to
 CC make a model for (I) or its complexes or parts, in X-ray crystallography
 CC techniques, or for determining three-dimensional structures of
 CC polypeptides with unknown structures. Pharmaceutical compositions
 CC comprising the ligand or modulator is useful for treating
 CC lymphoproliferative conditions, malignant and pre-malignant conditions
 CC (such as cancer), arthritis, inflammation, autoimmune disorder (such as
 CC lupus erythematosus, immune-related glomerulonephritis, rheumatoid
 CC arthritis), viral infection, inflammation, graft versus host disease,
 CC neurodegenerative diseases and conditions involving trauma and injury to
 CC the nervous system (e.g., Alzheimer's disease, Parkinson's disease,
 CC Huntington's disease and multiple sclerosis). The present sequence
 CC represents a VAB-1 protein sequence
 XX
 SQ Sequence 332 AA;

Query Match 82.8%; Score 24; DB 5; Length 332;
 Best Local Similarity 71.4%; Pred. No. 7.8e+02;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 VMXVAEF 7
 || || || ||
 Db 109 VMIVABY 115

RESULT 34
 AAU37753
 ID AAU37753 standard; protein; 335 AA.
 XX
 AC AAU37753;
 XX
 DT 14-FEB-2002 (first entry)
 XX
 DE Streptococcus pneumoniae cellular proliferation protein #182.
 XX
 KW Antisense; prokaryotic cellular proliferation protein; antibiotic;
 KW antibacterial; drug design.
 XX
 OS Streptococcus pneumoniae.
 XX
 PN WO200170955-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 21-MAR-2001; 2001WO-US009180.
 XX
 PR 21-MAR-2000; 2000US-0191078P.
 PR 23-MAY-2000; 2000US-0206848P.
 PR 26-MAY-2000; 2000US-0207727P.
 PR 23-OCT-2000; 2000US-0242578P.
 PR 27-NOV-2000; 2000US-0253625P.
 PR 22-DEC-2000; 2000US-0257931P.
 PR 16-FEB-2001; 2001US-0269308P.
 XX
 PA (ELIT-) ELITRA PHARM INC.
 XX
 PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
 PI Yamamoto RT, Xu HH;
 XX
 DR WPI; 2001-611495/70.
 DR N-PSDB; AAS55612.
 XX
 PT New polynucleotides for the identification and development of
 PT antibiotics, comprise sequences of antisense nucleic acids.
 XX
 PS Example 3; SEQ ID NO 13346; 511pp; English.
 XX
 CC The invention relates to antisense inhibitors of genes essential to
 CC prokaryotic cellular proliferation, their use in identifying the genes,
 CC their use in the discovery of novel antibiotics, the essential genes
 CC themselves and the encoded proteins. The prokaryotes used are Escherichia
 CC coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
 CC Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also

CC variant of the polynucleotide or a nucleic acid sequence 95% identical to
 CC one of the polynucleotides. The S. pneumoniae polynucleotides and encoded
 CC polypeptides (ABP81299-ABP81674) are useful for treating or preventing S.
 CC pneumoniae infections or non-systemic diseases, e.g. otitis media, which
 CC are induced or exacerbated by S. pneumoniae. These are also useful for
 CC detecting S. pneumoniae in a biological sample or diagnosing S.
 CC pneumoniae infection in a subject. The polynucleotides have antibacterial
 CC activity and are useful in gene therapy
 XX
 SQ Sequence 335 AA;

Query Match 82.8%; Score 24; DB 6; Length 335;
 Best Local Similarity 71.4%; Pred. No. 7.9e+02;
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
 |||||
 DB 224 VMSAAEF 230

RESULT 37
 ABU45990
 ID ABU45990 standard; protein; 335 AA.
 XX AC ABU45990;
 XX 19-JUN-2003 (first entry)
 DT
 DE Protein encoded by Prokaryotic essential gene #31517.
 XX
 XX Antisense; prokaryotic essential gene; cell proliferation; drug design.
 XX Streptococcus pneumoniae.
 XX WO200277183-A2.
 XX 03-OCT-2002.
 XX 21-MAR-2002; 2002WO-US009107.
 XX 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 XX (ELIT-) ELITRA PHARM INC.
 XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX WPI; 2003-029926/02.
 DR N-PSDB; ACA49860.
 XX New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX Claim 25; SEQ ID NO 73914; 1766pp; English.

CC The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)

CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than S. aureus, S. typhimurium,
 CC K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 335 AA;

Query Match 82.8%; Score 24; DB 6; Length 335;
 Best Local Similarity 71.4%; Pred. No. 7.9e+02;
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
 |||||
 DB 224 VMSAAEF 230

RESULT 38
 ADK48584
 ID ADK48584 standard; protein; 335 AA.
 XX AC ADK48584;
 XX 20-MAY-2004 (first entry)
 DT
 DE Streptococcus pneumoniae protein, Seq ID No 5099.
 XX
 XX Antibacterial; Gene therapy; Vaccine; Streptococcus pneumoniae.
 KW Streptococcus pneumoniae.
 OS US6699703-B1.
 PN
 XX 02-MAR-2004.
 PD
 XX 26-MAY-2000; 2000US-00583110.
 PF
 XX 02-JUL-1997; 97US-0051553P.
 PR 12-MAY-1998; 98US-0085131P.
 PR 30-JUN-1998; 98US-00107433.
 XX (GENO-) GENOME THERAPEUTICS CORP.
 XX Doucette-Stamm L, Bush D, Zeng Q, Opperman T, Househeart CE;
 PI WPI; 2004-212399/20.
 DR N-PSDB; ADK45923.
 XX New nucleic acid molecules and polypeptides useful for diagnosing,
 PT preventing and treating pathological conditions resulting from bacterial
 PT infection, e.g. Streptococcus pneumoniae infection, and in drug
 PT screening.
 XX Disclosure; SEQ ID NO 5099; 301pp; English.
 PS The invention relates to isolated Streptococcus pneumoniae nucleic acids
 CC and polypeptides. The nucleic acids and proteins are useful for
 CC diagnosing, preventing and treating pathological conditions resulting
 CC from bacterial infection, such as S. pneumoniae infection. These may also
 CC be used for drug screening procedures. The present sequence represents a

PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148311P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149358P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149829P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 82.8%; Score 24; DB 3; Length 361;
Best Local Similarity 71.4%; Pred. No. 8.5e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEP 7
Db 31 VMNVAEP 37
|||:||||

RESULT 40
AAG45883
ID AAG45883 standard; protein; 361 AA.
XX
AC AAG45883;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 57658.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.

PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147915P.
PR 10-AUG-1999; 99US-0148171P.

PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 82.8%; Score 24; DB 3; Length 361;
Best Local Similarity 71.4%; Pred. No. 8.5e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 VMXVAEF 7
DB 31 VMNMAEF 37

RESULT 41

AAG24458

ID AAG24458 standard; protein; 361 AA.

XX AAG24458;

AC AAG24458;

DT 17-OCT-2000 (first entry)

XX

XX

DE Arabidopsis thaliana protein fragment SEQ ID NO: 28136.

XX

KW Protein identification; signal transduction pathway; metabolic pathway;
hybridisation assay; genetic mapping; gene expression control; promoter;
termination sequence.

KW

XX

OS Arabidopsis thaliana.

XX

PN EP1033405-A2.

XX

PD 06-SEP-2000.

XX

PF 25-FEB-2000; 2000EP-00301439.

XX

PR 25-FEB-1999; 99US-0121825P.

PR

PR 05-MAR-1999; 99US-0123180P.

PR

PR 09-MAR-1999; 99US-0123548P.

PR

PR 21-MAR-1999; 99US-0125788P.

PR

PR 25-MAR-1999; 99US-0126264P.

PR

PR 29-MAR-1999; 99US-0126785P.

PR

PR 01-APR-1999; 99US-0127462P.

PR

PR 06-APR-1999; 99US-0128234P.

PR

PR 08-APR-1999; 99US-0128714P.

PR

PR 16-APR-1999; 99US-0129845P.

PR

PR 19-APR-1999; 99US-0130077P.

PR

PR 21-APR-1999; 99US-0130449P.

PR

PR 23-APR-1999; 99US-0130510P.

PR

PR 28-APR-1999; 99US-0130891P.

PR

PR 30-APR-1999; 99US-0132048P.

PR

PR 30-APR-1999; 99US-0132407P.

PR

PR 04-MAY-1999; 99US-0132484P.

PR

PR 05-MAY-1999; 99US-0132485P.

PR

PR 06-MAY-1999; 99US-0132486P.

PR

PR 06-MAY-1999; 99US-0132487P.

PR

PR 07-MAY-1999; 99US-0132863P.

PR

PR 11-MAY-1999; 99US-0134256P.

PR

PR 14-MAY-1999; 99US-0134218P.

PR

PR 14-MAY-1999; 99US-0134219P.

PR

PR 14-MAY-1999; 99US-0134221P.

PR

PR 18-MAY-1999; 99US-0134768P.

PR

PR 19-MAY-1999; 99US-0134941P.

PR

PR 20-MAY-1999; 99US-0135124P.

PR

PR 21-MAY-1999; 99US-0135353P.

PR

PR 24-MAY-1999; 99US-0135629P.

PR

PR 25-MAY-1999; 99US-0136021P.

PR

PR 27-MAY-1999; 99US-0136392P.

PR

PR 28-MAY-1999; 99US-0136782P.

PR

PR 01-JUN-1999; 99US-0137222P.

PR

PR 03-JUN-1999; 99US-0137528P.

PR

PR 04-JUN-1999; 99US-0137502P.

PR

PR 07-JUN-1999; 99US-0137724P.

PR

PR 08-JUN-1999; 99US-0138094P.

PR

PR 10-JUN-1999; 99US-0138540P.

PR

PR 10-JUN-1999; 99US-0138847P.

PR

PR 14-JUN-1999; 99US-0139119P.

PR

PR 16-JUN-1999; 99US-0139452P.

PR

PR 16-JUN-1999; 99US-0139453P.

PR

PR 17-JUN-1999; 99US-0139452P.

PR

PR 18-JUN-1999; 99US-0139454P.

PR

PR 18-JUN-1999; 99US-0139455P.

PR

PR 18-JUN-1999; 99US-0139456P.

PR

PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 28-JUN-1999; 99US-0140981P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147028P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.

PR	20-AUG-1999;	99US-0149722P.	XX	19-JUN-2003 (first entry)	
PR	20-AUG-1999;	99US-0149723P.	DT	Protein encoded by Prokaryotic essential gene #9133.	
PR	20-AUG-1999;	99US-0149923P.	XX	Antisense; prokaryotic essential gene; cell proliferation; drug design.	
PR	23-AUG-1999;	99US-0149920P.	DE	Clostridium acetobutylicum.	
PR	23-AUG-1999;	99US-0149930P.	XX	WO200277183-A2.	
PR	25-AUG-1999;	99US-0150566P.	XX	03-OCT-2002.	
PR	26-AUG-1999;	99US-0150884P.	XX	21-MAR-2002; 2002WO-US009107.	
PR	27-AUG-1999;	99US-0151065P.	XX	21-MAR-2001; 2001US-00815242.	
PR	27-AUG-1999;	99US-0151066P.	XX	06-SEP-2001; 2001US-00948993.	
PR	27-AUG-1999;	99US-0151080P.	PD	25-OCT-2001; 2001US-0342923P.	
PR	30-AUG-1999;	99US-0151130P.	XX	08-FEB-2002; 2002US-00072851.	
PR	31-AUG-1999;	99US-0151438P.	XX	06-MAR-2002; 2002US-0362699P.	
PR	01-SEP-1999;	99US-0151930P.	PF	(ELIT-) ELITRA PHARM INC.	
PR	07-SEP-1999;	99US-0152363P.	XX	Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen Ku, Zyskind JW;	
PR	10-SEP-1999;	99US-0153070P.	XX	Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;	
PR	13-SEP-1999;	99US-0153758P.	XX	WPI; 2003-029926/02.	
PR	15-SEP-1999;	99US-0154018P.	DR	N-PSDB; ACA27476.	
PR	16-SEP-1999;	99US-0154039P.	XX	New antisense nucleic acids, useful for identifying proteins or screening	
PR	20-SEP-1999;	99US-0154779P.	PT	for homologous nucleic acids required for cellular proliferation to	
PR	22-SEP-1999;	99US-0155139P.	PT	isolate candidate molecules for rational drug discovery programs.	
PR	23-SEP-1999;	99US-0155486P.	XX	Claim 25; SEQ ID NO 51530; 1766pp; English.	
PR	24-SEP-1999;	99US-0155659P.	PS	The invention relates to an isolated nucleic acid comprising any one of	
PR	28-SEP-1999;	99US-0156458P.	XX	the 6213 antisense sequences given in the specification where expression	
PR	29-SEP-1999;	99US-0156596P.	CC	of the nucleic acid inhibits proliferation of a cell. Also included are:	
PR	04-OCT-1999;	99US-0157117P.	CC	(1) a vector comprising a promoter operably linked to the nucleic acid	
PR	05-OCT-1999;	99US-0157753P.	CC	encoding a polypeptide whose expression is inhibited by the antisense	
PR	06-OCT-1999;	99US-0157865P.	CC	nucleic acid; (2) a host cell containing the vector; (3) an isolated	
PR	07-OCT-1999;	99US-0158029P.	CC	polypeptide or its fragment whose expression is inhibited by the	
PR	08-OCT-1999;	99US-0158232P.	CC	antisense nucleic acid; (4) an antibody capable of specifically binding	
PR	12-OCT-1999;	99US-0158369P.	CC	the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular	
PR	13-OCT-1999;	99US-0159253P.	CC	proliferation or the activity of a gene in an operon required for	
PR	13-OCT-1999;	99US-0159294P.	CC	proliferation; (7) identifying a compound that influences the activity of	
PR	13-OCT-1999;	99US-0159295P.	CC	the gene product or that has an activity against a biological pathway	
PR	14-OCT-1999;	99US-0159332P.	CC	required for proliferation, or that inhibits cellular proliferation; (8)	
PR	14-OCT-1999;	99US-0159333P.	CC	identifying a gene required for cellular proliferation or the biological	
PR	14-OCT-1999;	99US-0159637P.	CC	pathway in which a proliferation-required gene or its gene product lies	
PR	14-OCT-1999;	99US-0159638P.	CC	or a gene on which the test compound that inhibits proliferation of an	
PR	18-OCT-1999;	99US-0159584P.	CC	organism acts; (9) manufacturing an antibiotic; (10) profiling a	
PR	21-OCT-1999;	99US-0160741P.	CC	compound's activity; (11) a culture comprising strains in which the gene	
PR	21-OCT-1999;	99US-0160767P.	CC	product is overexpressed or underexpressed; (12) determining the extent	
PR	21-OCT-1999;	99US-0160768P.	CC	to which each of the strains is present in a culture or collection of	
PR	21-OCT-1999;	99US-0160814P.	CC	strains; or (13) identifying the target of a compound that inhibits the	
PR	21-OCT-1999;	99US-0160815P.	CC	proliferation of an organism. The antisense nucleic acids are useful for	
PR	22-OCT-1999;	99US-0160980P.	CC	identifying proteins or screening for homologous nucleic acids required	
PR	22-OCT-1999;	99US-0160981P.	CC	for cellular proliferation to isolate candidate molecules for rational	
PR	22-OCT-1999;	99US-0160989P.	CC	drug discovery programs, or for screening homologous nucleic acids	
PR	25-OCT-1999;	99US-0161404P.	CC	required for proliferation in cells other than <i>S. aureus</i> , <i>S. typhimurium</i> ,	
PR	25-OCT-1999;	99US-0161405P.	CC	<i>K. pneumoniae</i> or <i>P. aeruginosa</i> . The present sequence is encoded by one of	
PR	25-OCT-1999;	99US-0161359P.	CC	the target prokaryotic essential genes. Note: The sequence data for this	
PR	26-OCT-1999;	99US-0161360P.	CC	patent did not form part of the printed specification, but was obtained	
PR	26-OCT-1999;	99US-0161361P.	CC	in electronic format directly from WIPO at	
PR	28-OCT-1999;	99US-0161920P.	CC	ftp.wipo.int/pub/published_pct_sequences	
PR	28-OCT-1999;	99US-0161922P.	XX	Sequence 367 AA;	
PR	28-OCT-1999;	99US-0161933P.	SQ	Query Match 82.8%; Score 24; DB 6; Length 367;	
PR	29-OCT-1999;	99US-0162142P.		Best Local Similarity 71.4%; Pred. No. 8.7e+02;	
				Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;	
Qy	1 VMXVAEF 7				
	:				
Db	31 VMNVAEF 37				
RESULT 42					
ID	ABU23606				
XX	ABU23606 standard; protein; 367 AA.				
AC	ABU23606;				

Qy	1 VMXVABF 7	Best Local Similarity 71.4%; Pred. No. 8.8e+02;	
Db	29 VMEVABY 35	Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;	
RESULT 43			
AA	AAAM40200	AAAM40200 standard; protein; 373 AA.	
XX	AC	AAAM40200;	
XX	DE	22-OCT-2001 (first entry)	
XX	DE	Human polypeptide SEQ ID NO 3345.	
XX	KW	Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;	
XX	KW	peripheral nervous system; neuropathy; central nervous system; CNS;	
XX	KW	Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;	
XX	KW	amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;	
XX	KW	chemokinetic; thrombolytic; drug screening; arthritis; inflammation;	
XX	KW	leukaemia.	
XX	OS	Homo sapiens.	
XX	PN	WO200153312-A1.	
XX	PD	26-JUL-2001.	
XX	PF	26-DEC-2000; 2000WO-US034263.	
XX	PR	23-DEC-1999; 99US-00471275.	
XX	PR	21-JAN-2000; 2000US-00488725.	
XX	PR	25-APR-2000; 2000US-00552317.	
XX	PR	20-JUN-2000; 2000US-00598042.	
XX	PR	19-JUL-2000; 2000US-00620312.	
XX	PR	03-AUG-2000; 2000US-00653450.	
XX	PR	14-SEP-2000; 2000US-00662191.	
XX	PR	19-OCT-2000; 2000US-00693036.	
XX	PR	29-NOV-2000; 2000US-00727344.	
XX	PA	(HYSE-) HYSEQ INC.	
XX	PI	Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;	
XX	PI	Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;	
XX	PI	Zhou P, Goodrich R, Drmanac RT;	
XX	DR	WPI; 2001-442253/47.	
XX	DR	N-PSDB; AA159356.	
XX	PT	Novel nucleic acids and polypeptides, useful for treating disorders such	
XX	PT	as central nervous system injuries.	
XX	XX	Example 5; SEQ ID NO 3345; 10078pp; English.	
XX	PS	The invention relates to human nucleic acids (AA157798-AA161369) and the	
XX	CC	encoded polypeptides (AAAM38642-AAAM42213) with nootropic,	
XX	CC	immunosuppressant and cytostatic activity. The polynucleotides are useful	
XX	CC	in gene therapy. A composition containing a polypeptide or polynucleotide	
XX	CC	of the invention may be used to treat diseases of the peripheral nervous	
XX	CC	system, such as peripheral nervous injuries, peripheral neuropathy and	
XX	CC	localised neuropathies and central nervous system diseases, such as	
XX	CC	Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic	
XX	CC	lateral sclerosis, and Shy-Drager Syndrome. Other uses include the	
XX	CC	utilisation of the activities such as: immune system suppression,	
XX	CC	Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic	
XX	CC	and thrombolytic activity, cancer diagnosis and therapy, drug screening,	
XX	CC	assays for receptor activity, arthritis and inflammation, leukaemias and	
XX	CC	C.N.S disorders. Note: The sequence data for this patent did not form	
XX	CC	part of the printed specification	
XX	SQ	Sequence 373 AA;	
Query Match 82.8%; Score 24; DB 4; Length 373;			

Qy	1 VMXVABF 7	Best Local Similarity 71.4%; Pred. No. 8.8e+02;	
Db	305 LMVABF 311	Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;	
RESULT 44			
AA	AAAG20944	AAAG20944 standard; protein; 378 AA.	
XX	AC	AAAG20944;	
XX	DT	17-OCT-2000 (first entry)	
XX	DE	Arabidopsis thaliana protein fragment SEQ ID NO: 23320.	
XX	KW	Protein identification; signal transduction pathway; metabolic pathway;	
XX	KW	hybridisation assay; genetic mapping; gene expression control; promoter;	
XX	KW	termination sequence.	
XX	OS	Arabidopsis thaliana.	
XX	PN	EP1033405-A2.	
XX	PD	06-SEP-2000.	
XX	PF	25-FEB-2000; 2000EP-00301439.	
XX	PR	25-FEB-1999; 99US-0121825P.	
XX	PR	05-MAR-1999; 99US-0123180P.	
XX	PR	03-MAR-1999; 99US-0123548P.	
XX	PR	23-MAR-1999; 99US-0125788P.	
XX	PR	25-MAR-1999; 99US-0126264P.	
XX	PR	29-MAR-1999; 99US-0126785P.	
XX	PR	01-APR-1999; 99US-0127462P.	
XX	PR	06-APR-1999; 99US-0128234P.	
XX	PR	08-APR-1999; 99US-0128714P.	
XX	PR	16-APR-1999; 99US-0129845P.	
XX	PR	19-APR-1999; 99US-0130077P.	
XX	PR	21-APR-1999; 99US-0130449P.	
XX	PR	23-APR-1999; 99US-0130510P.	
XX	PR	23-APR-1999; 99US-0130891P.	
XX	PR	28-APR-1999; 99US-0131449P.	
XX	PR	30-APR-1999; 99US-0132048P.	
XX	PR	04-MAY-1999; 99US-0132407P.	
XX	PR	05-MAY-1999; 99US-0132484P.	
XX	PR	06-MAY-1999; 99US-0132485P.	
XX	PR	06-MAY-1999; 99US-0132486P.	
XX	PR	07-MAY-1999; 99US-0132487P.	
XX	PR	11-MAY-1999; 99US-0132863P.	
XX	PR	14-MAY-1999; 99US-0134218P.	
XX	PR	14-MAY-1999; 99US-0134219P.	
XX	PR	14-MAY-1999; 99US-0134221P.	
XX	PR	18-MAY-1999; 99US-0134370P.	
XX	PR	19-MAY-1999; 99US-0134768P.	
XX	PR	20-MAY-1999; 99US-0134941P.	
XX	PR	21-MAY-1999; 99US-0135124P.	
XX	PR	21-MAY-1999; 99US-0135353P.	
XX	PR	24-MAY-1999; 99US-0135629P.	
XX	PR	25-MAY-1999; 99US-0136021P.	
XX	PR	27-MAY-1999; 99US-0136392P.	
XX	PR	28-MAY-1999; 99US-0136782P.	
XX	PR	01-JUN-1999; 99US-0137222P.	
XX	PR	03-JUN-1999; 99US-0137528P.	
XX	PR	04-JUN-1999; 99US-0137502P.	
XX	PR	07-JUN-1999; 99US-0137724P.	
XX	PR	08-JUN-1999; 99US-0138094P.	
XX	PR	10-JUN-1999; 99US-0138540P.	
XX	PR	10-JUN-1999; 99US-0138847P.	
XX	PR	14-JUN-1999; 99US-0139119P.	

PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140895P.
PR 28-JUN-1999; 99US-0140823P.
PR 30-JUN-1999; 99US-0140991P.
PR 01-JUL-1999; 99US-01411287P.
PR 01-JUL-1999; 99US-0141842P.
PR 02-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 20-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.

PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 82.8%; Score 24; DB 3; Length 378;
Best Local Similarity 71.4%; Pred. No. 9e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7

Db 48 VMNMAEF 54

RESULT 45
AAG24457
ID AAG24457 standard; protein; 378 AA.
XX
AC AAG24457;
XX
DT 17-OCT-2000 (first entry)
XX
XX DE - Arabidopsis thaliana protein fragment SEQ ID NO: 28135.
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
XX Arabidopsis thaliana.
XX
PN EF1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 08-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 21-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 08-JUL-1999; 99US-0142820P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144684P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.

PR	20-AUG-1999;	99US-0149723P.	DT	18-OCT-2000	(first entry)
PR	20-AUG-1999;	99US-0149829P.	XX		
PR	23-AUG-1999;	99US-0149902P.	DE	Arabidopsis thaliana protein fragment	SEQ ID NO: 57657.
PR	23-AUG-1999;	99US-0149930P.	XX		
PR	25-AUG-1999;	99US-0150566P.	KW	Protein identification; signal transduction pathway; metabolic pathway;	
PR	26-AUG-1999;	99US-0150884P.	KW	Hybridisation assay; genetic mapping; gene expression control; promoter;	
PR	27-AUG-1999;	99US-0151065P.	KW	termination sequence.	
PR	27-AUG-1999;	99US-0151066P.	XX		
PR	27-AUG-1999;	99US-0151080P.	OS	Arabidopsis thaliana.	
PR	30-AUG-1999;	99US-0151303P.	XX		
PR	31-AUG-1999;	99US-0151438P.	PN	EP1033405-A2.	
PR	01-SEP-1999;	99US-0151930P.	XX		
PR	07-SEP-1999;	99US-0152363P.	PD	06-SEP-2000.	
PR	10-SEP-1999;	99US-0153070P.	XX		
PR	13-SEP-1999;	99US-0153758P.	PF	25-FEB-2000;	2000EP-00301439.
PR	15-SEP-1999;	99US-0154018P.	XX		
PR	16-SEP-1999;	99US-0154039P.	PR	25-FEB-1999;	99US-0121825P.
PR	20-SEP-1999;	99US-0154779P.	PR	05-MAR-1999;	99US-0123180P.
PR	22-SEP-1999;	99US-0155139P.	PR	09-MAR-1999;	99US-0123548P.
PR	23-SEP-1999;	99US-0155486P.	PR	23-MAR-1999;	99US-0125788P.
PR	24-SEP-1999;	99US-0155859P.	PR	25-MAR-1999;	99US-0126264P.
PR	28-SEP-1999;	99US-0156458P.	PR	29-MAR-1999;	99US-0126785P.
PR	29-SEP-1999;	99US-0156596P.	PR	01-APR-1999;	99US-0127462P.
PR	04-OCT-1999;	99US-0157117P.	PR	06-APR-1999;	99US-0128234P.
PR	05-OCT-1999;	99US-0157753P.	PR	08-APR-1999;	99US-0128714P.
PR	06-OCT-1999;	99US-0157865P.	PR	16-APR-1999;	99US-0129845P.
PR	07-OCT-1999;	99US-0158029P.	PR	19-APR-1999;	99US-0130077P.
PR	08-OCT-1999;	99US-0158232P.	PR	21-APR-1999;	99US-0130449P.
PR	12-OCT-1999;	99US-0158369P.	PR	23-APR-1999;	99US-0130510P.
PR	13-OCT-1999;	99US-0159293P.	PR	23-APR-1999;	99US-0130891P.
PR	13-OCT-1999;	99US-0159294P.	PR	28-APR-1999;	99US-0131449P.
PR	13-OCT-1999;	99US-0159295P.	PR	30-APR-1999;	99US-0132048P.
PR	14-OCT-1999;	99US-0159329P.	PR	30-APR-1999;	99US-0132407P.
PR	14-OCT-1999;	99US-0159330P.	PR	04-MAY-1999;	99US-0132484P.
PR	14-OCT-1999;	99US-0159331P.	PR	05-MAY-1999;	99US-0132485P.
PR	14-OCT-1999;	99US-0159637P.	PR	06-MAY-1999;	99US-0132486P.
PR	14-OCT-1999;	99US-0159638P.	PR	06-MAY-1999;	99US-0132487P.
PR	18-OCT-1999;	99US-0159584P.	PR	07-MAY-1999;	99US-0132863P.
PR	21-OCT-1999;	99US-0160741P.	PR	11-MAY-1999;	99US-0134256P.
PR	21-OCT-1999;	99US-0160767P.	PR	14-MAY-1999;	99US-0134218P.
PR	21-OCT-1999;	99US-0160768P.	PR	14-MAY-1999;	99US-0134219P.
PR	21-OCT-1999;	99US-0160770P.	PR	14-MAY-1999;	99US-0134221P.
PR	21-OCT-1999;	99US-0160814P.	PR	14-MAY-1999;	99US-0134370P.
PR	21-OCT-1999;	99US-0160815P.	PR	18-MAY-1999;	99US-0134768P.
PR	22-OCT-1999;	99US-0160860P.	PR	19-MAY-1999;	99US-0134941P.
PR	22-OCT-1999;	99US-0160989P.	PR	20-MAY-1999;	99US-0135124P.
PR	22-OCT-1999;	99US-0160989P.	PR	21-MAY-1999;	99US-0135353P.
PR	25-OCT-1999;	99US-0161404P.	PR	24-MAY-1999;	99US-0135629P.
PR	25-OCT-1999;	99US-0161405P.	PR	25-MAY-1999;	99US-0136021P.
PR	26-OCT-1999;	99US-0161359P.	PR	27-MAY-1999;	99US-0136392P.
PR	26-OCT-1999;	99US-0161360P.	PR	28-MAY-1999;	99US-0136782P.
PR	26-OCT-1999;	99US-0161361P.	PR	01-JUN-1999;	99US-0137222P.
PR	26-OCT-1999;	99US-0161361P.	PR	03-JUN-1999;	99US-0137528P.
PR	28-OCT-1999;	99US-0161920P.	PR	04-JUN-1999;	99US-0137502P.
PR	28-OCT-1999;	99US-0161992P.	PR	07-JUN-1999;	99US-0137724P.
PR	28-OCT-1999;	99US-0161993P.	PR	08-JUN-1999;	99US-0138094P.
PR	28-OCT-1999;	99US-0162142P.	PR	10-JUN-1999;	99US-0138540P.
PR	28-OCT-1999;	99US-0162142P.	PR	10-JUN-1999;	99US-0138847P.
PR	28-OCT-1999;	99US-0162142P.	PR	14-JUN-1999;	99US-0139119P.
PR	28-OCT-1999;	99US-0162142P.	PR	16-JUN-1999;	99US-0139452P.
PR	28-OCT-1999;	99US-0162142P.	PR	16-JUN-1999;	99US-0139453P.
PR	28-OCT-1999;	99US-0162142P.	PR	17-JUN-1999;	99US-0139492P.
PR	28-OCT-1999;	99US-0162142P.	PR	18-JUN-1999;	99US-0139454P.
PR	28-OCT-1999;	99US-0162142P.	PR	18-JUN-1999;	99US-0139455P.
PR	28-OCT-1999;	99US-0162142P.	PR	18-JUN-1999;	99US-0139456P.
PR	28-OCT-1999;	99US-0162142P.	PR	18-JUN-1999;	99US-0139457P.
PR	28-OCT-1999;	99US-0162142P.	PR	18-JUN-1999;	99US-0139458P.
PR	28-OCT-1999;	99US-0162142P.	PR	18-JUN-1999;	99US-0139459P.
PR	28-OCT-1999;	99US-0162142P.	PR	18-JUN-1999;	99US-0139460P.
PR	28-OCT-1999;	99US-0162142P.	PR	18-JUN-1999;	99US-0139461P.
PR	28-OCT-1999;	99US-0162142P.	PR	18-JUN-1999;	99US-0139462P.
PR	28-OCT-1999;	99US-0162142P.	PR	18-JUN-1999;	99US-0139463P.
PR	28-OCT-1999;	99US-0162142P.	PR	18-JUN-1999;	99US-0139750P.

Query Match 82.8%; Score 24; DB 3; Length 378;
Best Local Similarity 71.4%; Pred. No. 9e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 VMKVAEF 7
Db 48 VMNNAEF 54

RESULT 46
AAG45882
ID AAG45882 standard; protein; 378 AA.
XX
AC AAG45882;
XX

```
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 28-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 06-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
```

```
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161320P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.
```

Query Match 82.8%; Score 24; DB 3; Length 378;

Best Local Similarity 71.4%; Pred. No. 9e+02; Mismatches 1; Indels 0; Gaps 0;

Matches 5; Conservative 1;

QY 1 VMXVAEF 7

Db 48 VMXVAEF 54

RESULT 47

AAAY71056

ID AAAY71056 standard; protein; 384 AA.

XX AAAY71056;

AC AAAY71056;

XX 29-AUG-2000 (first entry)

DT Human membrane transport protein, MTRP-1.

XX Human; membrane transport protein; MTRP-1; antiinflammatory; cytostatic;

XX antithyroid; immunosuppressive; thyromimetic; antidiabetic; nootropic;

XX antidiarrheic; neuroprotective; antidepressant; nephrotropic; virucide;

KW Human; membrane transport protein; MTRP-1; antiinflammatory; cytostatic;

KW antithyroid; immunosuppressive; thyromimetic; antidiabetic; nootropic;

KW antidiarrheic; neuroprotective; antidepressant; nephrotropic; virucide;

CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: Immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemias and
 CC C.N.S disorders. Note: The sequence data for this patent did not form
 CC part of the printed specification
 XX
 XX Sequence 384 AA;

Query Match 82.8%; Score 24; DB 4; Length 384;
 Best Local Similarity 71.4%; Pred. No. 9.1e+02;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VMXVAEF 7
 Db 305 LMVVAEF 311

RESULT 49
 AAM40199
 ID AAM40199 standard; protein; 384 AA.
 AC AAM40199;
 XX

DT 22-OCT-2001 (first entry)

DE Human polypeptide SEQ ID NO 3344.

XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 KW leukaemia.

XX Homo sapiens.

XX WO200153312-A1.

XX 26-JUL-2001.

XX 26-DEC-2000; 2000WO-US034263.

XX 23-DEC-1999; 99US-00471275.

XX 21-JAN-2000; 2000US-00488725.

XX 25-APR-2000; 2000US-00552317.

XX 20-JUN-2000; 2000US-00598042.

XX 19-JUL-2000; 2000US-00620312.

XX 03-AUG-2000; 2000US-00653450.

XX 14-SEP-2000; 2000US-00662191.

XX 19-OCT-2000; 2000US-00693036.

XX 29-NOV-2000; 2000US-00727344.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
 PI Zhou P, Goodrich R, Drmanac RT;

XX WPI; 2001-442253/47.

XX N-PSDB; AAI59355.

XX Novel nucleic acids and polypeptides, useful for treating disorders such
 PT as central nervous system injuries.

XX Example 5; SEQ ID NO 3344; 10078pp; English.

XX The invention relates to human nucleic acids (AAI57798-AAI61369) and the
 CC encoded polypeptides (AAM38642-AAM42213) with nootropic,

CC immunosuppressant and cytostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: Immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemias and
 CC C.N.S disorders. Note: The sequence data for this patent did not form
 CC part of the printed specification
 XX

SQ Sequence 384 AA;

Query Match 82.8%; Score 24; DB 4; Length 384;
 Best Local Similarity 71.4%; Pred. No. 9.1e+02;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VMXVAEF 7
 Db 305 LMVVAEF 311

RESULT 50

AAE30494

ID AAE30494 standard; protein; 409 AA.

XX AAE30494;

DT 24-FEB-2003 (first entry)

DE Fruit fly gustatory receptor protein, Gr64A3.

XX Fruit fly; gustatory receptor; Gr; odorant receptor; Or; pesticide;

XX Gr64A3.

XX Drosophila melanogaster.

XX WO200268593-A2.

XX 06-SEP-2002.

XX 22-FEB-2002; 2002WO-US005414.

XX 23-FEB-2001; 2001US-0271319P.

XX (UYCO) UNIV COLUMBIA NEW YORK.

XX Axel R, Scott K;

XX WPI; 2002-698668/75.

XX Novel nucleic acid encoding insect gustatory or odorant receptor protein
 PT useful for identifying a compound which specifically binds to the
 PT receptor for controlling a pest population in an area.

XX Claim 3; Page 164-166; 264pp; English.

XX The invention relates to an isolated nucleic acid encoding an insect
 CC gustatory receptor (Gr) or odorant receptor (Or) protein which comprises
 CC seven transmembrane domains and a C-terminal domain comprising
 CC consecutive amino acids. The invention is useful for identifying a
 CC compound which activates the insect receptor or inhibits the activity of
 CC the insect receptor. The purified insect receptor protein is embedded in
 CC a lipid bilayer. The invention is sprayed for combating ingestion of
 CC crops by pest insects, combating disease-carrying insects in an area and
 CC controlling a pest population in an area. The invention is useful for
 CC detecting the presence of insect gustatory or odorant receptor and for
 CC inhibiting the function of the receptor in humans or animals or in
 CC biological fluids isolated from them. The invention is also useful for
 CC identifying or isolating other insect receptors and for combating pest

```
CC nuisances. The present sequence is fruit fly gustatory receptor (Gr)
CC protein
XX
SQ Sequence 409 AA;

Query Match      82.8%; Score 24; DB 5; Length 409;
Best Local Similarity 71.4%; Pred. No. 9.8e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy      1 VMXVAEF 7
      | : ||||
Db      24 VLLVAEF 30

RESULT 51
AAG4581
ID AAG45881 standard; protein; 414 AA.
XX
AC AAG45881;
DX
DT 18-OCT-2000 (first entry)
XX
XX Arabidopsis thaliana protein fragment SEQ ID NO: 57656.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
PN EP1033405-A2.
PD 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-00301439.
XX
PR 26-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 02-JUL-1999; 99US-0142154P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
```



```
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 23-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156488P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159233P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 18-OCT-1999; 99US-0159638P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161933P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 82.8%; Score 24; DB 3; Length 414;
Best Local Similarity 71.4%; Pred. No. 9.9e+02;
```

```
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 VMXVAEF 7
Db 84 VMXVAEF 90

RESULT 52
AAG24456
ID AAG24456 standard; protein; 414 AA.
XX
AC AAG24456;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 28134.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-00301439.
XX
XX 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 25-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 04-MAY-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 07-MAY-1999; 99US-0132487P.
PR 11-MAY-1999; 99US-0132863P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 18-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 04-JUN-1999; 99US-0137528P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
```

PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 02-JUL-1999; 99US-0142154P.
PR 08-JUL-1999; 99US-0142055P.
PR 08-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 28-JUL-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147301P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.

PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 27-AUG-1999; 99US-0151080P.
PR 31-AUG-1999; 99US-0151303P.
PR 01-SEP-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 82.8%; Score 24; DB 3; Length 414;
Best Local Similarity 71.4%; Pred. No. 9.9e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
||:||||
Db 84 VMMAEF 90

RESULT 53

AAG20943
ID AAG20943 standard; protein; 414 AA.
XX
AC AAG20943;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 23319.
XX
DE Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 16-APR-1999; 99US-0128714P.
PR 19-APR-1999; 99US-0129845P.
PR 21-APR-1999; 99US-0130077P.
PR 23-APR-1999; 99US-0130449P.
PR 28-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131149P.
PR 30-APR-1999; 99US-0132048P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 07-MAY-1999; 99US-0132487P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 18-MAY-1999; 99US-0134370P.
PR 19-MAY-1999; 99US-0134768P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 21-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 02-JUL-1999; 99US-0142154P.
PR 06-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145911P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.

```
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 25-AUG-1999; 99US-0149930P.
PR 26-AUG-1999; 99US-0150566P.
PR 27-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 24-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161360P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 82.8%; Score 24; DB 3; Length 414;
Best Local Similarity 71.4%; Pred. No. 9.9e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
DB 84 VMNNAEF 90

RESULT 54
ADRO9407
XX ADRO9407 standard; protein; 457 AA.
XX ADR09407;
XX ADR09407;
DT 04-NOV-2004 (first entry)

XX 99US-0149929P.
DE 99US-0149902P.
XX 99US-0149930P.
KW 99US-0150566P.
KW 99US-0150884P.
KW 99US-0151065P.
KW 99US-0151066P.
KW 99US-0151080P.
KW 99US-0151303P.
XX 99US-0151438P.
OS 99US-0151930P.
XX 99US-0152363P.
PN 99US-0153070P.
XX 99US-0153758P.
PD 99US-0154018P.
XX 99US-0154039P.
PF 99US-0154779P.
XX 99US-0155139P.
PR 99US-0155486P.
PR 99US-0155659P.
XX 99US-0156458P.
PA (REAS-) RES ASSOC BIOTECHNOLOGY.
XX 99US-0157117P.
PI 99US-0157753P.
PI 99US-0157865P.
XX 99US-0158029P.
DR 99US-0158232P.
DR N-PSDB; ADR07451.
XX 99US-0158369P.
PT 99US-0159293P.
PT 99US-0159294P.
XX 99US-0159295P.
PS 99US-0159329P.
XX 99US-0159330P.
XX 99US-0159331P.
CC This invention relates to novel, isolated full length human cDNA
CC molecules and the encoded proteins thereof. Specifically, it refers to
CC cDNA clones obtained by an oligo-capping method, where none of these
CC clones are identical to any known human mRNAs. The present invention
CC describes an immunoassay to identify agonists and antagonists, as well as
CC antibodies, antisense molecules and siRNAs that can all be used to bind
CC to and modulate expression of the cDNA molecules. As such, these
CC molecules are useful for diagnostic markers or therapeutic targets for
CC the various diseases or morbid states. In particular, they are useful in
CC gene therapy for treating osteoporosis, neurological disease, Alzheimer's
CC disease, Parkinson's disease, dementia, short memory and various cancers,
CC as well as for maintaining equilibrium of sense or motor function, and
CC for treating emotional reaction, fear response and panic. Accordingly,
CC they exhibit osteopathic, neuroprotective, nootropic, antiparkinsonian,
CC cytotstatic and tranquiliser activities. This polypeptide is a protein
CC encoded by a full length human cDNA sequence of the invention. NOTE: This
CC sequence is not given in the sequence listing of the specification but
CC can be obtained on CD-ROM from the European Patent Office, Vienna Sub-
CC office.
XX Sequence 457 AA;
SQ

Query Match 82.8%; Score 24; DB 8; Length 457;
Best Local Similarity 71.4%; Pred. No. 1.1e+03;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
DB 389 LMVVAEF 395

RESULT 55
AAB96296
ID AAB96296 standard; protein; 584 AA.
XX AAB96296;
AC AAB96296;
XX AAB96296;
DT 29-OCT-2001 (first entry)
XX
```



```

PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
XX 23-MAR-2001; 2001WO-US009231.
XX PF
XX 23-MAR-2000; 2000US-0191637P.
PR
PR 11-JUL-2000; 2000US-00614150.
XX
XX (PEKE ) PE CORP NY.
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
XX
XX WPI; 2001-656860/75.
DR N-PSDB; ABL14130.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.
XX
XX Disclosure; SEQ ID NO 36873; 2lpp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 817 AA;
SQ
Query Match 82.8%; Score 24; DB 4; Length 817;
Best Local Similarity 71.4%; Pred. No. 2.1e+03;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VMXVAEF 7
Db 432 VLLVAEF 438
|: |||||

RESULT 58
ABU18144
ID ABU18144 standard; protein; 902 AA.
XX
XX ABU18144;
XX
XX 19-JUN-2003 (first entry)
XX
XX Protein encoded by Prokaryotic essential gene #3671.
XX
XX Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
XX Bacillus anthracis.
XX
XX WO200277183-A2.
XX
XX 03-OCT-2002.
XX
XX 21-MAR-2002; 2002WO-US009107.
XX
XX 21-MAR-2001; 2001US-00815242.
PR
PR 06-SEP-2001; 2001US-00948993.
PR
PR 25-OCT-2001; 2001US-0342923P.
PR
PR 08-FEB-2002; 2002US-00072851.
PR
PR 08-MAR-2002; 2002US-0362659P.
XX
XX (ELIT-) ELITRA PHARM INC.
XX
XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI

```

```

PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
DR WPI; 2003-029926/02.
DR N-PSDB; ACA22014.
XX
XX New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
XX Claim 25; SEQ ID NO 46069; 1766pp; English.
PS
XX
XX The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than S. aureus, S. typhimurium,
CC K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 902 AA;
SQ
Query Match 82.8%; Score 24; DB 6; Length 902;
Best Local Similarity 83.3%; Pred. No. 2.3e+03;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 MXVAEF 7
Db 113 MTVAEF 118
|: |||||

RESULT 59
ABU42813
ID ABU42813 standard; protein; 919 AA.
XX
XX ABU42813;
XX
XX 19-JUN-2003 (first entry)
XX
XX Protein encoded by Prokaryotic essential gene #28340.
XX
XX Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
XX Staphylococcus epidermidis.
XX
XX WO200277183-A2.
XX
XX 03-OCT-2002.
XX
XX 21-MAR-2002; 2002WO-US009107.
XX

```

XX 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 XX (ELIT-) ELITRA PHARM INC.
 XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX WPI; 2003-0299926/02..
 DR N-PSDB; ACA46683.
 XX New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX Claim 25; SEQ ID NO 70737; 1766pp; English.
 XX The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC product is overexpressed or underexpressed, (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs; or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX SQ Sequence 919 AA;
 Query Match 82.8%; Score 24; DB 6; Length 919;
 Best Local Similarity 83.3%; Pred. No. 2.3e+03;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 2 MXVAEF 7
 Db 118 MSVAEF 123
 RESULT 60
 ADN22714
 ID ADN22714 standard; protein; 919 AA.
 XX AC ADN22714;
 XX 02-DEC-2004 (first entry)
 DT Bacterial polypeptide #5367.
 XX

XX Recombinant DNA construct; transformed plant; improved plant property;
 KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
 KW pathogen tolerance; pest tolerance; plant disease resistance;
 KW cell cycle pathway modification; plant growth regulator;
 KW homologous recombination; seed oil yield; protein yield; carbohydrate;
 KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
 KW bacterial polypeptide.
 XX Bacteria.
 XX US2003233675-A1.
 XX 18-DEC-2003.
 XX 20-FEB-2003; 2003US-00369493.
 XX 21-FEB-2002; 2002US-0360039P.
 XX (CAOV/) CAO Y.
 PA (HINK/) HINKLE G J.
 PA (SLAT/) SLATER S C.
 PA (CHEN/) CHEN X.
 XX (GOLD/) GOLDMAN B S.
 PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
 XX WPI; 2004-061375/06.
 DR New recombinant DNA construct comprising a promoter positioned to provide
 PT for expression of a polynucleotide encoding a polypeptide from a
 PT microbial source, useful for producing plants with improved properties.
 XX Claim 1; SEQ ID NO 5367; 122pp; English.
 XX The invention relates to a recombinant DNA construct comprising a
 CC promoter functional in a plant cell, where the promoter is positioned to
 CC provide for expression of a polynucleotide encoding a polypeptide from a
 CC microbial source. The invention also relates to a transformed plant
 CC comprising the recombinant DNA construct and a method of producing a
 CC transformed plant having an improved property. The plant is a crop plant
 CC such as maize or soybean. The method of producing a transformed plant
 CC having an improved property comprises transforming a plant with the
 CC recombinant DNA construct and growing the transformed plant, where the
 CC polynucleotide or polypeptide is useful for improving plant properties.
 CC The recombinant DNA construct is useful for producing plants with
 CC improved plant properties, e.g. improved cold, heat or drought tolerance,
 CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
 CC increased resistance to plant disease, better growth rate by modification
 CC of the cell cycle pathway with plant growth regulators, increased rate of
 CC homologous recombination, modified seed oil or protein yield and/or
 CC content, improved yield by modification of carbohydrate, nitrogen or
 CC phosphorus use and/or uptake, by modification of photosynthesis or by
 CC providing improved plant growth and development under at least one stress
 CC condition, improved lignin production or improved galactomannan
 CC production. This sequence represents a bacterial polypeptide used in the
 CC scope of the invention. Note: The sequence data for this patent did not
 CC form part of the printed specification but was obtained in electronic
 CC format from USPTO at seqdata.uspto.gov/sequence.html.
 XX SQ Sequence 919 AA;
 Query Match 82.8%; Score 24; DB 8; Length 919;
 Best Local Similarity 71.4%; Pred. No. 2.3e+03;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 VMXVAEF 7
 Db 592 VMIVAEY 598
 RESULT 61
 ADB11880

```

ID ADB11880 standard; protein; 926 AA.
XX
AC ADB11880;
XX
DT 20-NOV-2003 (first entry)
XX
DE Alloicoccus otitis antigenic protein SEQ ID NO:5180.
XX
KW Alloicoccus otitidis; antigenic protein; immunogenic; immunisation;
XX gene therapy; Gram-positive bacterium; infection.
XX
OS Alloicoccus otitis.
XX
PN W02003048304-A2.
XX
PD 12-JUN-2003.
XX
PF 25-NOV-2002; 2002WO-US036123.
XX
PR 29-NOV-2001; 2001US-0333777P.
XX
PR 18-NOV-2002; 2002US-0426742P.
XX
PA (AMHP ) WYETH HOLDINGS CORP.
XX
PI Fletcher LD, McMichael JC, Russell DP, Zagursky RJ;
XX
XX WPI; 2003-505284/47.
XX
DR N-PSDB; ADB11883.
XX
XX New Alloicoccus otitidis polynucleotides and polypeptides, useful for
XX treating and diagnosing diseases, drug screening assays and monitoring of
XX effects during drug clinical trials.
XX
PS Claim 33; SEQ ID NO 5180; 1019pp; English.
XX
XX The present invention describes an isolated polynucleotide (I) of
XX Alloicoccus otitidis genomic DNA, which encodes an antigenic protein.
XX Alloicoccus otitidis is a Gram-positive bacterium. Also described: (1)
XX an isolated polypeptide that is encoded by the polynucleotide (I); (2) an
XX expression vector comprising the novel isolated polynucleotide (I); its
XX complement, degenerate variant or fragment; (3) a genetically engineered
XX host cell, transfected, transformed or infected with the vector of (2);
XX (4) an antibody specific for the polypeptide of (1); (5) an immunogenic
XX composition comprising the polypeptide, its complement, biological
XX equivalent or fragment, or the polynucleotide that is comprised in the
XX expression vector; (6) a pharmaceutical composition comprising the
XX polypeptide of (1) and a carrier; (7) a protein chip comprising an array
XX of the polypeptides of (1), their biological equivalent or fragment; (8)
XX immunising against Alloicoccus otitidis by administering to a host the
XX immunogenic composition; (9) detecting and/or identifying Alloicoccus
XX otitidis in the biological sample; (10) a kit comprising a container
XX containing the novel polynucleotide, its degenerate variant or fragment,
XX or the antibody of (4); and (11) producing a polypeptide by culturing the
XX genetically engineered host cell under conditions suitable to produce the
XX polypeptide from the culture. (1) can be used in gene therapy. The
XX polynucleotides, polypeptides, antibodies and compositions of the present
XX invention can be used for treating and diagnosing diseases, drug
XX screening assays and monitoring of effects during drug clinical trials.
XX The polynucleotides are useful for expressing and detecting Alloicoccus
XX otitidis. The present sequence represents an Alloicoccus otitidis
XX antigen protein from the present invention.
XX
SQ Sequence 926 AA;
Query Match 82.8%; Score 24; DB 6; Length 926;
Best Local Similarity 83.3%; Pred. No. 2.4e+03;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 MXVAEF 7
Db 115 MSVAEF 120
| | | | |

RESULT 63
ID ADS06396
AD ADS06396 standard; protein; 927 AA.
XX
AC ADS06396;
XX
DT 04-NOV-2004 (first entry)
XX
DE Staphylococcus epidermis polypeptide seqid 5691.
XX
KW antibacterial; vaccine; antisense therapy; Staphylococcus epidermidis;
XX recombinant expression vector; infection; computer readable medium;
XX computer based system.
XX
OS Staphylococcus epidermidis.

```

```

RESULT 62
ABP39986
ID ABP39986 standard; protein; 927 AA.
XX
AC ABP39986;
XX
DT 24-JUL-2002 (first entry)
XX
DE Staphylococcus epidermidis ORF amino acid sequence SEQ ID NO:4831.
XX
KW Staphylococcus epidermidis; open reading frame; ORF; bacterial infection;
XX antibacterial; gene therapy.
XX
OS Staphylococcus epidermidis.
XX
PN US6380370-B1.
XX
PD 30-APR-2002.
XX
PF 13-AUG-1998; 98US-00134001.
XX
PR 14-AUG-1997; 97US-0055779P.
XX
PR 08-NOV-1997; 97US-0064964P.
XX
PA (GENO-) GENOME THERAPEUTICS CORP.
XX
PI Doucette-Stamm LA, Bush D;
XX
XX WPI; 2002-381255/41.
XX
DR N-PSDB; ABN92531.
XX
XX Novel isolated nucleic acid encoding a Staphylococcus epidermis
XX polypeptide, useful for diagnosing and treating bacterial infections.
XX
PS Disclosure; SEQ ID NO 4831; 267pp; English.
XX
XX ABN90538 to ABN93374 represent Staphylococcus epidermidis open reading
XX frame (ORF) nucleic acid sequences which encode the amino acid sequences
XX given in ABP35124 to ABP37960. The S. epidermidis sequences have
XX antibacterial activity and can be used in gene therapy. The sequences can
XX also be used in the diagnosis and treatment of bacterial infections,
XX particularly S. epidermidis infections. The sequences can be used to
XX screen for compounds able to interfere with the S. epidermidis life cycle
XX or inhibit S. epidermidis infection. N.B. The sequence data for this
XX patent did not form part of the printed specification, but was obtained
XX in electronic format directly from the USPTO web site
XX
SQ Sequence 927 AA;
Query Match 82.8%; Score 24; DB 5; Length 927;
Best Local Similarity 83.3%; Pred. No. 2.4e+03;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 MXVAEF 7
Db 126 MSVAEF 131
| | | | |

RESULT 63
ID ADS06396
AD ADS06396 standard; protein; 927 AA.
XX
AC ADS06396;
XX
DT 04-NOV-2004 (first entry)
XX
DE Staphylococcus epidermis polypeptide seqid 5691.
XX
KW antibacterial; vaccine; antisense therapy; Staphylococcus epidermidis;
XX recombinant expression vector; infection; computer readable medium;
XX computer based system.
XX
OS Staphylococcus epidermidis.

```


XX US2004147734-A1.
 XX 29-JUL-2004.
 XX
 XX 01-DEC-2003; 2003US-00724972.
 XX
 XX 08-NOV-1997; 97US-0064964P.
 XX 13-AUG-1998; 98US-00134001.
 XX 29-NOV-1999; 99US-00450969.
 XX
 XX (DOUC/) DOUCETTE-STAMM L.
 XX (BUSH/) BUSH D.
 XX
 XX Doucette-Stamm L, Bush D;
 XX
 XX WPI; 2004-580138/56.
 XX N-PSDB; ADS02624.
 XX
 XX New isolated polypeptide and encoding nucleic acid derived from
 PT Staphylococcus epidermidis, useful for diagnosing, preventing and/or
 PT treating an S. epidermidis bacterial infection.
 XX
 XX Claim 17; SEQ ID NO 5691; 741pp; English.
 XX
 CC The invention describes an isolated nucleic acid comprising a nucleotide
 CC sequence with any of 3772 fully defined nucleotide sequences (SEQ ID NO:
 CC 1-3772) and encoding a Staphylococcus epidermidis polypeptide with any
 CC of 3772 fully defined amino acid sequences (SEQ ID NO: 3772-7544) as
 CC given in the specification. Also described are: a recombinant expression
 CC vector; a cell comprising a recombinant expression vector of (1);
 CC producing an S. epidermidis polypeptide; an isolated nucleic acid
 CC comprising a nucleotide sequence of at least 8 nucleotides in length; a
 CC vaccine composition for prevention or treatment of an S. epidermidis
 CC infection, comprising a nucleic acid cited above and a carrier; treating
 CC a subject for S. epidermidis infection; a recombinant or substantially
 CC pure preparation of an S. epidermidis polypeptide or its fragment; a
 CC vaccine composition for prevention or treatment of an S. epidermidis
 CC infection; detecting the presence of a Staphylococcus nucleic acid in a
 CC sample; a computer readable medium having recorded in it the nucleotide
 CC sequences with SEQ ID NO: 1-3772 or its fragments; a computer based
 CC system for identifying fragments of the Staphylococcus genome of
 CC commercial importance; a computer based system for identifying fragments
 CC of the Staphylococcus plasmids of commercial importance; identifying
 CC commercially important nucleic acid fragments of the Staphylococcus
 CC genome and/or plasmids; and identifying an expression modulating fragment
 CC of the Staphylococcus genome and/or plasmids. The methods and
 CC compositions of the present invention are useful for the diagnosis,
 CC prevention and/or treatment of an Staphylococcal epidermidis bacterial
 CC infection. This is the amino acid sequence of a S. epidermidis protein of
 CC the invention.
 XX
 SQ Sequence 927 AA;
 Query Match 82.8%; Score 24; DB 8; Length 927;
 Best Local Similarity 83.3%; Pred. No. 2.4e+03;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 2 MXVAEF 7
 Db 126 MSVAEF 131
 RESULT 64
 ADB11878
 ID ADB11878 standard; protein; 938 AA.
 XX AC ADB11878;
 XX
 XX 20-NOV-2003 (first entry)
 XX
 XX Alloococcus otitis antigenic protein SEQ ID NO:5182.
 XX
 KW Alloococcus otitidis; antigenic protein; immunogenic; immunisation;
 KW gene therapy; Gram-positive bacterium; infection.
 XX
 OS Alloococcus otitis.
 XX
 PN WO2003048304-A2.
 XX
 XX 12-JUN-2003.
 XX
 XX 25-NOV-2002; 2002WO-US036123.
 XX
 XX 29-NOV-2001; 2001US-0333777P.
 XX 18-NOV-2002; 2002US-0426742P.
 PR
 XX (AMHP) WYETH HOLDINGS CORP.
 XX
 XX Fletcher LD, Mcmichael JC, Russell DP, Zagursky RJ;
 PI
 XX WPI; 2003-505284/47.
 DR N-PSDB; ADB11881.
 DR
 XX New Alloococcus otitidis polynucleotides and polypeptides, useful for
 PT treating and diagnosing diseases, drug screening assays and monitoring of
 PT effects during drug clinical trials.
 XX
 XX Claim 33; SEQ ID NO 5182; 1019pp; English.
 XX
 CC The present invention describes an isolated polynucleotide (I) of
 CC Alloococcus otitidis genomic DNA, which encodes an antigenic protein.
 CC Alloococcus otitidis is a Gram-positive bacterium. Also described: (1)
 CC an isolated polypeptide that is encoded by the polynucleotide (I); (2) an
 CC expression vector comprising the novel isolated polynucleotide (I), its
 CC complement, degenerate variant or fragment; (3) a genetically engineered
 CC host cell, transfected, transformed or infected with the vector of (2);
 CC (4) an antibody specific for the polypeptide of (1); (5) an immunogenic
 CC composition comprising the polypeptide, its complement, biological
 CC equivalent or fragment, or the polynucleotide that is comprised in the
 CC expression vector; (6) a pharmaceutical composition comprising the
 CC polypeptide of (1) and a carrier; (7) a protein chip comprising an array
 CC of the polypeptides of (1), their biological equivalent or fragment; (8)
 CC immunising against Alloococcus otitidis by administering to a host the
 CC immunogenic composition; (9) detecting and/or identifying Alloococcus
 CC otitidis in the biological sample; (10) a kit comprising a container
 CC containing the novel polynucleotide, its degenerate variant or fragment,
 CC or the antibody of (4); and (11) producing a polypeptide by culturing the
 CC genetically engineered host cell under conditions suitable to produce the
 CC polypeptide from the culture. (i) can be used in gene therapy. The
 CC polynucleotides, polypeptides, antibodies and compositions of the present
 CC invention can be used for treating and diagnosing diseases, drug
 CC screening assays and monitoring of effects during drug clinical trials.
 CC The polynucleotides are useful for expressing and detecting Alloococcus
 CC otitidis. The present sequence represents an Alloococcus otitidis
 CC antigen protein from the present invention.
 XX
 SQ Sequence 938 AA;
 Query Match 82.8%; Score 24; DB 6; Length 938;
 Best Local Similarity 83.3%; Pred. No. 2.4e+03;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 2 MXVAEF 7
 Db 127 MSVAEF 132
 RESULT 65
 ADB11876
 ID ADB11876 standard; protein; 940 AA.
 XX AC ADB11876;
 XX
 XX 20-NOV-2003 (first entry)
 XX
 XX Alloococcus otitis antigenic protein SEQ ID NO:5182.
 XX

DE Alloiooccus otitis antigenic protein SEQ ID NO:5184.
 XX
 KW Alloiooccus otitidis; antigenic protein; immunogenic; immunisation;
 KW gene therapy; Gram-positive bacterium; infection.
 XX
 OS Alloiooccus otitis.
 XX
 PN WO2003048304-A2.
 XX
 PD 12-JUN-2003.
 XX
 XX 25-NOV-2002; 2002WO-US036123.
 XX
 PF 29-NOV-2001; 2001US-0333777P.
 PR 18-NOV-2002; 2002US-0426742P.
 XX
 XX (AMHP) WYETH HOLDINGS CORP.
 PA
 XX Fletcher LD, McMichael JC, Russell DP, Zagursky RJ;
 PI
 DR WPI; 2003-505284/47.
 DR N-PSDB; ADB11879.
 XX
 XX New Alloiooccus otitidis polynucleotides and polypeptides, useful for
 PT treating and diagnosing diseases, drug screening assays and monitoring of
 PT effects during drug clinical trials.
 PT
 XX
 PS Claim 33; SEQ ID NO 5184; 1019pp; English.
 XX
 CC The present invention describes an isolated polynucleotide (I) of
 CC Alloiooccus otitidis genomic DNA, which encodes an antigenic protein.
 CC Alloiooccus otitidis is a Gram-positive bacterium. Also described: (1)
 CC an isolated polypeptide that is encoded by the polynucleotide (I); (2) an
 CC expression vector comprising the novel isolated polynucleotide (I), its
 CC complement, degenerate variant or fragment; (3) a genetically engineered
 CC host cell, transfected, transformed or infected with the vector of (2);
 CC (4) an antibody specific for the polypeptide of (1); (5) an immunogenic
 CC composition comprising the polypeptide, its complement, biological
 CC equivalent or fragment, or the polynucleotide that is comprised in the
 CC expression vector; (6) a pharmaceutical composition comprising the
 CC polypeptide of (1) and a carrier; (7) a protein chip comprising an array
 CC of the polypeptides of (1), their biological equivalent or fragment; (8)
 CC immunising against Alloiooccus otitidis by administering to a host the
 CC immunogenic composition; (9) detecting and/or identifying Alloiooccus
 CC otitidis in the biological sample; (10) a kit comprising a container
 CC containing the novel polynucleotide, its degenerate variant or fragment,
 CC or the antibody of (4); and (11) producing a polypeptide by culturing the
 CC genetically engineered host cell under conditions suitable to produce the
 CC polypeptide from the culture. (I) can be used in gene therapy. The
 CC polynucleotides, polypeptides, antibodies and compositions of the present
 CC invention can be used for treating and diagnosing diseases, drug
 CC screening assays and monitoring of effects during drug clinical trials.
 CC The polynucleotides are useful for expressing and detecting Alloiooccus
 CC otitidis. The present sequence represents an Alloiooccus otitidis
 CC antigen protein from the present invention.
 XX
 SQ Sequence 940 AA;
 Query Match 82.8%; Score 24; DB 6; Length 940;
 Best Local Similarity 83.3%; Pred. No. 2.4e+03;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 2 MXVAEF 7
 | ||||
 Db 129 MSVAEF 134
 RESULT 66
 AAR58701
 ID AAR58701 standard; protein; 962 AA.
 XX
 AC AAR58701;
 XX

DT 25-MAR-2003 (revised)
 DT 25-MAR-1995 (first entry)
 XX
 DE Aminopectidase H110D helminthic antigen.
 XX
 KW Aminopectidase; H110D; antigen; vaccine.
 XX
 OS Haemonchus contortus.
 XX
 PN WO9418320-A1.
 XX
 PD 18-AUG-1994.
 XX
 PF 04-FEB-1994; 94WO-GB000204.
 XX
 PR 05-FEB-1993; 93GB-00002302.
 XX
 PA (MLCW) MALLINCKRODT VETERINARY INC.
 PA (AGRI-) AGRIC & FOOD RES COUNCIL.
 XX
 PI Munn E, Graham M, Smith TS, Rolph TP, Newton SE;
 XX
 DR WPI; 1994-279741/34.
 DR N-PSDB; AAQ71611.
 XX
 PT Expression of parasite gut membrane bound antigen - for protective
 PT immunity against parasites esp. Haemonchus contortus.
 XX
 PS Disclosure; Page 23-27; 37pp; English.
 XX
 CC The Haemonchus contortus enzyme antigen H110D is an aminopeptidase, and
 CC is expressed in COS1 or CHO cells. The recombinant antigen is used in a
 CC vaccine for the protection of mammals and animals against the parasite H.
 CC contortus. (Updated on 25-MAR-2003 to correct PN field.)
 XX
 SQ Sequence 962 AA;
 Query Match 82.8%; Score 24; DB 2; Length 962;
 Best Local Similarity 71.4%; Pred. No. 2.5e+03;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 VMXVAEF 7
 |: ||||
 Db 532 VISVAEF 538
 RESULT 67
 AAR51280
 ID AAR51280 standard; protein; 972 AA.
 XX
 AC AAR51280;
 XX
 XX 25-MAR-2003 (revised)
 DT 31-MAY-1994 (first entry)
 XX
 DE Helminth aminopeptidase H110D variant encoded by the H11-3 gene.
 XX
 KW Helminth; aminopeptidase; H11-3; H11-2; H11-1; integral membrane;
 KW alpha-amino acyl peptide hydrolase (microsomal); multigene family;
 KW antigen; vaccine; parasite; human; virus; microbe.
 XX
 OS Haemonchus contortus.
 XX
 PN WO93232542-A1.
 XX
 PD 25-NOV-1993.
 XX
 PF 07-MAY-1993; 93WO-GB0000943.
 XX
 PR 08-MAY-1992; 92GB-00009993.
 XX
 PA (AGRI-) AGRIC & FOOD RES COUNCIL.
 XX

PI Graham M, Smith TS, Munn EA, Knox DP, Oliver JJ, Newton SE;
 XX WPI; 1993-386574/48.
 DR N-PSDB; AAQ52489.
 XX
 PT New DNA encoding amino-peptidase from *Haemonchus contortus* and derived
 PT peptide(s) - useful in protective vaccines, for new vectors, transformed
 PT cells and oligosaccharide, for incorporation in virus or microbe.
 XX
 PS Claim 6; Page 82-84; 137pp; English.
 XX
 CC The sequences given in AAR51280-82 represent variants of the helminth
 CC aminopeptidase H10D encoded by the genes H11-3, -2 and -1 respectively.
 CC H10D is a protein doublet which shows homology to a family of integral
 CC membrane aminopeptidases. The differences between the coding sequences
 CC and the translated products, can be attributed to different mRNAs of the
 CC multigene family, and also to different variants of the H10D-encoding
 CC sequence being present at different stages of the life cycle, or in
 CC strains differing in geographical origin. Antigenic fragments of these
 CC aminopeptidases may be used in vaccines to stimulate immune response
 CC against helminth parasites in humans or other animals. The DNA sequences
 CC encoding these proteins may be incorporated into a virus or microbe and
 CC used in a similar manner. (Updated on 25-MAR-2003 to correct PN field.)
 XX
 SQ - Sequence 972 AA;
 Query Match 82.8%; Score 24; DB 2; Length 972;
 Best Local Similarity 71.4%; Pred. No. 2.5e+03;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 VMXVAEF 7
 : : : : :
 Db 542 VISVAEF 548
 RESULT 68
 ABP52161
 ID ABP52161 standard; protein; 1084 AA.
 XX
 AC ABP52161;
 XX
 DT 14-OCT-2002 (first entry)
 XX
 DE Human 67084FL transporter protein SEQ ID NO:23.
 XX
 KW Human; transporter protein; anorectic; antidiabetic; anti-Parkinsonian;
 KW nootropic; neuroprotective; hypotensive; antidepressant; neuroleptic;
 KW cardiovascular; immunosuppressive; gene therapy; sugar homeostasis;
 KW obesity; diabetes; anorexia; central nervous system disorder; depression;
 KW CNS disorder; Parkinson's disease; Alzheimer's disease; hypertension;
 KW autonomic function disorder; schizophrenia; learning disorder; amnesia;
 KW memory disorder; age-related disorder; cardiovascular disorder;
 KW ischaemia reperfusion injury; restenosis; hormonal disorder;
 KW hypothyroidism; hyperthyroidism; immune disorder;
 KW chronic mucocutaneous candidiasis.
 XX
 OS Homo sapiens.
 XX
 PN WO200255701-A2.
 XX
 PD 18-JUL-2002.
 XX
 PF 17-DEC-2001; 2001WO-US049060.
 XX
 PR 15-DEC-2000; 2000US-0256240P.
 PR 18-DEC-2000; 2000US-0256588P.
 PR 21-DEC-2000; 2000US-0258028P.
 XX
 PA (MILL-) MILLENNIUM PHARM INC.
 XX
 PI Curtis RAJ;
 XX
 DR WPI; 2002-590672/63.

DR N-PSDB; ABQ74274, ABQ74275.
 XX
 PT New 67076, 67102, 44181, 67084FL and 67084alt nucleic acids and proteins,
 PT useful in e.g. treating disorders characterized by insufficient or
 PT excessive production of the polypeptides, e.g. cardiovascular or immune
 PT disorders.
 XX
 PS Claim 13; Fig 32; 364pp; English.
 XX
 CC The present sequence represents the human 67084FL protein, which is a
 CC member of the transporter family. The transporter family proteins (I) of
 CC the present invention have anorectic, antidiabetic, anti-Parkinsonian,
 CC nootropic, neuroprotective, hypotensive, antidepressant, neuroleptic,
 CC cardiovascular and immunosuppressive activities, and can be used in gene
 CC therapy. The human transporter proteins from the present invention are
 CC designated 8099, 46455, 54414, 53763, 67076, 67102, 44181, 67084FL and
 CC 67084alt. The polynucleotides encoding (I) can be used in screening
 CC assays (e.g. chromosome mapping, tissue typing, or in forensic biology),
 CC predictive medicine (e.g. diagnostic assays, prognostic assays,
 CC monitoring clinical trials, or pharmacogenetics), as surrogate markers,
 CC and in methods of treatment (e.g. therapeutic or prophylactic). (I) are
 CC useful for treating disorders characterised by insufficient or excessive
 CC production of 8099, 46455, 54414, 53763, 67076, 67102, 44181, 67084FL or
 CC 67084alt substrates or production of transport 8099, 46455, 54414, 53763,
 CC with 8099 and 46455 include disorders associated with sugar homeostasis
 CC such as obesity, diabetes or anorexia. Disorders associated with 54414,
 CC 53763, 67076, 67102, 44181, 67084FL or 67084alt are e.g. CNS disorders
 CC (e.g. Parkinson's disease or Alzheimer's disease), autonomic function
 CC disorders (e.g. hypertension, depression or schizophrenia), or learning
 CC or memory disorders (e.g. amnesia or age-related disorders),
 CC cardiovascular disorders (e.g. ischaemia reperfusion injury or
 CC restenosis), hormonal disorders (e.g. hypothyroidism or hyperthyroidism),
 CC or immune disorders (e.g. chronic mucocutaneous candidiasis)
 XX
 SQ Sequence 1084 AA;
 Query Match 82.8%; Score 24; DB 5; Length 1084;
 Best Local Similarity 71.4%; Pred. No. 2.8e+03;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 VMXVAEF 7
 : : : : :
 Db 1016 LMVVAEF 1022
 RESULT 69
 ABP52175
 ID ABP52175 standard; protein; 1084 AA.
 XX
 AC ABP52175;
 XX
 DT 14-OCT-2002 (first entry)
 XX
 DE Human 67084alt related transporter protein SEQ ID NO:27.
 XX
 KW Human; transporter protein; anorectic; antidiabetic; anti-Parkinsonian;
 KW nootropic; neuroprotective; hypotensive; antidepressant; neuroleptic;
 KW cardiovascular; immunosuppressive; gene therapy; sugar homeostasis;
 KW obesity; diabetes; anorexia; central nervous system disorder; depression;
 KW CNS disorder; Parkinson's disease; Alzheimer's disease; hypertension;
 KW autonomic function disorder; schizophrenia; learning disorder; amnesia;
 KW memory disorder; age-related disorder; cardiovascular disorder;
 KW ischaemia reperfusion injury; restenosis; hormonal disorder;
 KW hypothyroidism; hyperthyroidism; immune disorder;
 KW chronic mucocutaneous candidiasis.
 XX
 OS Homo sapiens.
 XX
 PN WO200255701-A2.
 XX
 PD 18-JUL-2002.

```

PF 17-DEC-2001; 2001WO-US049060.
XX
PR 15-DEC-2000; 2000US-0256240P.
PR 18-DEC-2000; 2000US-0256598P.
PR 21-DEC-2000; 2000US-0258028P.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
XX Curtis RAJ;
PI
XX
XX WPI; 2002-590672/63.
XX N-PSDB; ABQ74277.
XX
XX New 67076, 67102, 44181, 67084FL and 67084alt nucleic acids and proteins,
PT useful in e.g. treating disorders characterized by insufficient or
PT excessive production of the polypeptides, e.g. cardiovascular or immune
PT disorders.
XX
XX Disclosure; Page 334-339; 364pp; English.
XX
XX The present invention describes human proteins which are members of the
CC transporter family. The transporter family proteins (I) of the present
CC invention have anorectic, antidiabetic, anti-Parkinsonian, nootropic,
CC neuroprotective, hypotensive, antidepressant, neuroleptic, cardiovascular
CC and immunosuppressive activities, and can be used in gene therapy. The
CC human transporter proteins from the present invention are designated
CC 8099, 46455, 54414, 53763, 67076, 67102, 44181, 67084FL and 67084alt. The
CC polynucleotides encoding (I) can be used in screening assays (e.g.
CC chromosome mapping, tissue typing, or in forensic biology), predictive
CC medicine (e.g. diagnostic assays, prognostic assays, monitoring clinical
CC trials, or pharmacogenetics), as surrogate markers, and in methods of
CC treatment (e.g. therapeutic or prophylactic). (I) are useful for treating
CC disorders characterised by insufficient or excessive production of 8099,
CC 46455, 54414, 53763, 67076, 67102, 44181, 67084FL or 67084alt substrates
CC or production of transport 8099, 46455, 54414, 53763, 67076, 67102,
CC 44181, 67084FL or 67084alt inhibitors. Disorders associated with 8099 and
CC 46455 include disorders associated with sugar homeostasis such as
CC obesity, diabetes or anorexia. Disorders associated with 54414, 53763,
CC 67076, 67102, 44181, 67084FL or 67084alt are e.g. CNS disorders (e.g.
CC Parkinson's disease or Alzheimer's disease), autonomic function disorders
CC (e.g. hypertension, depression or schizophrenia), or learning or memory
CC disorders (e.g. amnesia or age-related disorders), cardiovascular
CC disorders (e.g. ischaemia reperfusion injury or restenosis), hormonal
CC disorders (e.g. hypothyroidism or hyperthyroidism), or immune disorders
CC (e.g. chronic mucocutaneous candidiasis). The present sequence represents
CC a human transporter protein related protein from the present invention
XX
XX Sequence 1084 AA;
SQ
Query Match 82.8%; Score 24; DB 5; Length 1084;
Best Local Similarity 71.4%; Pred. No. 2.8e+03;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 VMXVAEF 7
Db 1016 LMVVAEF 1022
:| |||
RESULT 70
ADD37496
ID ADD37496 standard; protein; 1084 AA.
XX
XX ADD37496;
AC
XX
XX 15-JAN-2004 (first entry)
DT
XX Human transporter 67084FL.
DE
XX Human; transporter; cytostatic; anorectic; antidiabetic; anticonvulsant;
KW gene therapy; PGC-1 associated disorder; liver tumour; obesity; epilepsy;
KW diabetes.
XX
XX Homo sapiens.
OS

```

```

XX
PN US2003143675-A1.
XX
XX 31-JUL-2003.
PD
XX
XX 22-MAY-2002; 2002US-00154419.
PF
XX
XX 12-MAY-2000; 2000US-0204211P.
PR 29-JUN-2000; 2000US-0215376P.
PR 31-JUL-2000; 2000US-0221769P.
PR 19-SEP-2000; 2000US-0233790P.
PR 25-SEP-2000; 2000US-0235107P.
PR 05-OCT-2000; 2000US-0238336P.
PR 14-NOV-2000; 2000US-0248364P.
PR 15-NOV-2000; 2000US-0248878P.
PR 18-DEC-2000; 2000US-0256240P.
PR 21-DEC-2000; 2000US-0256588P.
PR 22-JAN-2001; 2001US-0263169P.
PR 14-MAY-2001; 2001US-00858194.
PR 29-JUN-2001; 2001US-00895811.
PR 31-JUL-2001; 2001US-00919781.
PR 19-SEP-2001; 2001US-00957664.
PR 25-SEP-2001; 2001US-00964295.
PR 05-OCT-2001; 2001US-00972724.
PR 14-NOV-2001; 2001US-00002769.
PR 17-DEC-2001; 2001US-00024623.
PR 22-JAN-2002; 2002US-00055025.
XX
XX (MILL-) MILLENNIUM PHARM INC.
PA
XX Curtis RAJ, Gluckemann MA, Meyers RE;
PI
XX WPI; 2003-851783/79.
XX N-PSDB; ADD37495, ADD37497.
DR
XX
XX New isolated nucleic acid, useful for preparing a composition for
PT treating PGC-1 associated disorders e.g. liver tumors, obesity, epilepsy
PT or diabetes.
XX
XX Claim 11; SEQ ID NO 73; 663pp; English.
XX
XX The invention relates to an isolated nucleic acid comprising a cDNA
CC encoding a human transporter protein, or its complement, a sequence that
CC is 60 % identical to the cDNA, a fragment comprising at least 30
CC nucleotides of the cDNA, or a sequence encoding a fragment of the
CC polypeptide comprising at least 10 contiguous amino acid residues of the
CC cDNA. Also included are a vector comprising the novel nucleic acid
CC molecule, producing the polypeptide, the isolated transporter
CC polypeptide, an isolated antibody that specifically binds to the
CC polypeptide, detecting the presence of the polypeptide or nucleic acid in
CC a sample, a kit, identifying a compound that binds to, or that modulates
CC the activity of, the polypeptide, and modulating the activity of the
CC polypeptide. The nucleic acid is useful for preparing a composition for
CC treating PGC-1 (not defined) associated disorders e.g. liver tumours,
CC obesity, epilepsy or diabetes. The present sequence represents a novel
CC human transporter protein.
XX
XX Sequence 1084 AA;
SQ
Query Match 82.8%; Score 24; DB 7; Length 1084;
Best Local Similarity 71.4%; Pred. No. 2.8e+03;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 VMXVAEF 7
Db 1016 LMVVAEF 1022
:| |||
RESULT 71
AD127975
ID AD127975 standard; protein; 1084 AA.
XX

```

AD127975;
06-MAY-2004 (first entry)
Human 67084FL protein.
Human; 67084FL; ion channel family; ICF; cancer; leukaemia;
Alzheimer's disease; Parkinson's disease; multiple sclerosis; epilepsy;
hepatic disorder; cardiovascular disorder; cytostatic; neuroprotective;
nootropic; antiparkinsonian; hepatotropic; cardiovascular.
Homo sapiens.
US2003165891-A1.
04-SEP-2003.
15-MAY-2002; 2002US-00146733.
29-FEB-2000; 2000US-00515520.
29-FEB-2000; 2000US-0185938P.
03-MAR-2000; 2000US-00518866.
07-APR-2000; 2000US-0195734P.
11-APR-2000; 2000US-0195993P.
26-APR-2000; 2000US-0199799P.
19-SEP-2000; 2000US-0233537P.
25-SEP-2000; 2000US-0235018P.
25-SEP-2000; 2000US-0235059P.
15-DEC-2000; 2000US-0256240P.
18-DEC-2000; 2000US-0256588P.
21-DEC-2000; 2000US-0258028P.
28-FEB-2001; 2001US-00796720.
06-APR-2001; 2001US-00828035.
11-APR-2001; 2001US-00833081.
25-APR-2001; 2001US-00843128.
19-SEP-2001; 2001US-00957683.
25-SEP-2001; 2001US-00964252.
25-SEP-2001; 2001US-00964256.
17-DEC-2001; 2001US-00024623.
(MILL-) MILLENNIUM PHARM INC.
Curtis RAJ, Gluckemann MA, Silos-Santiago I;
WPI: 2004-069000/07.
N-PSDB; AD127974, AD127976.
TWIK-6, TWIK-7, IC23927, TWIK-8, IC47611, IC47615, HNMDA-1, TWIK-9,
alpha2delta-4, 54414, or 53763 nucleic acids and proteins, useful for
preventing, diagnosing and treating e.g. cancers, Alzheimer's disease and
cardiovascular disorders.
Example 1; SEQ ID NO 68; 638pp; English.
The invention relates to TWIK-6, TWIK-7, IC23927, TWIK-8, IC47611,
IC47615, HNMDA-1, TWIK-9, alpha-2delta-4, 54414 and 53763 ion channel
family (ICF) nucleic acids and proteins. The TWIK-6, TWIK-7, IC23927,
TWIK-8, IC47611, IC47615, HNMDA-1, TWIK-9, alpha-2delta-4, 54414 and
53763 ICF nucleic acids and proteins may be used for preventing,
diagnosing and treating ICF-related diseases. The sequences may be used
to treat disorders associated with decreased expression by rectifying
mutations or deletions in a patient's genome that affect the activity of
ICF proteins by expressing inactive proteins or to supplement the
patients own production of ICF proteins. The proteins may also be used as
antigens in the production of antibodies against ICF proteins and in
assays to identify modulators of ICF protein expression and activity. The
anti-ICF protein antibodies, agonists and antagonists may be used to
regulate ICF protein expression and activity. The antibodies may also be
used as diagnostic agents for detecting the presence of ICF proteins in
samples (e.g. by immunoassay). The nucleic acids and proteins may be used
to prevent, diagnose and treat a wide variety of disorders, e.g. cancers
and leukaemia, Alzheimer's disease, Parkinson's disease, multiple
sclerosis, epilepsy, hepatic disorders and cardiovascular disorders. This

CC sequence represents the human 67084FL protein of the invention. Note: The
CC sequence data for this patent is also available in electronic format from
CC USPTO at seqdata.uspto.gov/sequence.html.
XX
XX Sequence 1084 AA;
Query Match 82.8%; Score 24; DB 8; Length 1084;
Best Local Similarity 71.4%; Pred. No. 2.8e+03;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Qy 1 VMXVAEF 7
Db 1016 LMTVAEF 1022
RESULT 72
ADN33134
ID ADN33134 standard; protein; 1084 AA.
XX AC ADN33134;
XX 18-NOV-2004 (first entry)
XX Human transporter and ion channel (TRICH) protein SeqID19.
XX transporters and ion channel; TRICH; neuroprotective; relaxant;
XX antithyroid; antidiabetic; cytostatic; dermatological; immunosuppressive;
XX antiinflammatory; thyromimetic; antiallergic; cerebroprotective;
XX gastrointestinal; hepatotropic; nephrotropic; anticonvulsant;
XX antiparkinsonian; antibacterial; antiparasitic; fungicide; protozoicide;
XX virucide; uropathic; antirheumatic; cardiant; cardiovascular; anti-HIV;
XX nootropic; TRICH agonist; TRICH antagonist; gene therapy;
XX neurodegenerative disorder; Parkinson's disease; Alzheimer's disease;
XX muscular disorder; myotonic dystrophy; catatonias; endocrine disorder;
XX diabetes; Grave's disease; cancer; leukaemia; cervical cancer;
XX breast cancer; immunological disorder; scleroderma;
XX systemic lupus erythematosus; allergy; gastrointestinal disorder;
XX Crohn's disease; renal disorder; Goodpasture's syndrome; infection;
XX viral; bacterial; fungal; parasitic; protozoal; helminthic;
XX cardiovascular disorder; atherosclerosis; hepatic disease; cirrhosis;
XX transgenic animal; human.
XX Homo sapiens.
XX OS
XX WO2004035755-A2.
XX 29-APR-2004.
XX 16-OCT-2003; 2003WO-US033087.
XX 16-OCT-2002; 2002US-0419313P.
XX 23-OCT-2002; 2002US-0421033P.
XX 25-OCT-2002; 2002US-0421349P.
XX 04-NOV-2002; 2002US-0423516P.
XX (INCY-) INCYTE CORP.
XX Hafalia AJA, Khare R, Lal PG, Yue H, Baughn MR, Thornton MB;
XX Lu DAM, Ison CH, Becha SD, Ding L, Warren BA, Lee SY, Swarnakar A;
XX Elliott VS, Richardson TW, Marquis JP, Runkumar J, Murage J;
XX Rammann BE, Yao MG, Lu Y, Gietzen KJ, Yang YG, Chang H, Chawla NK;
XX Tran UK, Lee S, Yang J, Gandhi AR, Tribouley CM, Policky JL;
XX Ring HZ, Lee EA;
XX WPI: 2004-348448/32.
XX N-PSDB; ADN33130.
XX New TRICH polypeptides, useful for diagnosing, preventing, and treating
XX disorders associated with abnormal expression or activity of TRICH, e.g.
XX neuromuscular, immunological, cardiovascular disorders, cancer and/or
XX infections.
XX Claim 1; SEQ ID NO 19; 285pp; English.

XX This invention relates to novel human transporters and ion channel
 CC (TRICH) proteins and the nucleotide sequences which encode them. The
 CC invention may be useful for the production of compounds with a
 CC neuroprotective, relaxant, antihypertensive, antidiabetic, cytostatic,
 CC dermatological, immunosuppressive, antiinflammatory, thyromimetic,
 CC antiallergic, cerebroprotective, gastrointestinal, hepatotropic,
 CC nephrotropic, anticonvulsant, antiparkinsonian, antibacterial,
 CC antiparasitic, fungicide, protozoacide, virucide, uropathic,
 CC antitumour, cardiant, cardiovascular, anti-HIV or neurotropic activity
 CC acting as TRICH agonists or antagonists. In addition the disclosed
 CC sequences may be useful for gene therapy. The invention may be useful in
 CC diagnosing, preventing, and treating disorders associated with an
 CC abnormal expression or activity of TRICH, such as neurodegenerative
 CC disorders (for example Parkinson's disease, Alzheimer's disease),
 CC muscular disorders (for example myotonic dystrophy, catatonias), endocrine
 CC disorders (for example diabetes, Grave's disease), cancers (for example
 CC leukaemia, cervical or breast cancers), immunological disorders (for
 CC example scleroderma, systemic lupus erythematosus, allergies),
 CC gastrointestinal disorders (for example Crohn's disease), renal disorders
 CC (for example Goodpasture's syndrome), infections (for example viral,
 CC bacterial, fungal, parasitic, protozoal, helminthic), cardiovascular
 CC disorders (for example atherosclerosis), or hepatic diseases (for example
 CC cirrhosis). TRICH or its fragments may also be used in screening for
 CC compounds that specifically bind to and modulate the activity of TRICH.
 CC The nucleotides can be used to create humanised animals or transgenic
 CC animals to model human disease. The present sequence is that of a human
 CC transporter and ion channel (TRICH) protein of the invention.
 XX
 SQ Sequence 1084 AA;
 Query Match 82.8%; Score 24; DB 8; Length 1084;
 Best Local Similarity 71.4%; Pred. No. 2.8e+03;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 VMXVAEF 7
 :| |||||
 Db 1016 LMVVAEF 1022
 RESULT 73
 ABP52162
 ID ABP52162 standard; protein; 1095 AA.
 XX
 AC ABP52162;
 XX
 DT 14-OCT-2002 (first entry)
 XX
 DE Human 67084alt transporter protein SEQ ID NO:26.
 XX
 KW Human; transporter protein; anorectic; antidiabetic; anti-Parkinsonian;
 KW neurotropic; neuroprotective; hypotensive; antidepressant; neuroleptic;
 KW cardiovascular; immunosuppressive; gene therapy; sugar homeostasis;
 KW obesity; diabetes; anorexia; central nervous system disorder; depression;
 KW CNS disorder; Parkinson's disease; Alzheimer's disease; hypertension;
 KW autonomic function disorder; schizophrenia; learning disorder; amnesia;
 KW memory disorder; age-related disorder; cardiovascular disorder;
 KW ischaemia reperfusion injury; restenosis; hormonal disorder;
 KW hypothyroidism; hyperthyroidism; immune disorder;
 KW chronic mucocutaneous candidiasis.
 XX
 OS Homo sapiens.
 XX
 PN WO200255701-A2.
 XX
 XX 18-JUL-2002.
 XX
 PF 17-DEC-2001; 2001WO-US049060.
 XX
 PR 15-DEC-2000; 2000US-0256240P.
 PR 18-DEC-2000; 2000US-0256588P.
 PR 21-DEC-2000; 2000US-0258028P.
 XX

PA (MILL-) MILLENNIUM PHARM INC.
 XX
 PI Curtis RAJ;
 XX
 DR WPI; 2002-590672/63.
 DR N-PSDB; ABQ74276.
 XX
 PT New 67076, 67102, 44181, 67084FL and 67084alt nucleic acids and proteins,
 PT useful in e.g. treating disorders characterized by insufficient or
 PT excessive production of the polypeptides, e.g. cardiovascular or immune
 PT disorders.
 XX
 PS Claim 13; Fig 36; 364pp; English.
 XX
 CC The present sequence represents the human 67084alt protein, which is a
 CC member of the transporter family. The transporter family proteins (I) of
 CC the present invention have anorectic, antidiabetic, anti-Parkinsonian,
 CC neurotropic, neuroprotective, hypotensive, antidepressant, neuroleptic,
 CC cardiovascular and immunosuppressive activities, and can be used in gene
 CC therapy. The human transporter proteins from the present invention are
 CC designated 8099, 46455, 54414, 53763, 67076, 67102, 44181, 67084FL and
 CC 67084alt. The polynucleotides encoding (I) can be used in screening
 CC assays (e.g. chromosome mapping, tissue typing, or in forensic biology),
 CC predictive medicine (e.g. diagnostic assays, prognostic markers,
 CC monitoring clinical trials, or pharmacogenetics), as surrogate markers,
 CC and in methods of treatment (e.g. therapeutic or prophylactic). (I) are
 CC useful for treating disorders characterised by insufficient or excessive
 CC production of 8099, 46455, 54414, 53763, 67076, 67102, 44181, 67084FL or
 CC 67084alt substrates or production of transport 8099, 46455, 54414, 53763,
 CC 67076, 67102, 44181, 67084FL or 67084alt inhibitors. Disorders associated
 CC with 8099 and 46455 include disorders associated with sugar homeostasis
 CC such as obesity, diabetes or anorexia. Disorders associated with 54414,
 CC 53763, 67076, 67102, 44181, 67084FL or 67084alt are e.g. CNS disorders
 CC (e.g. Parkinson's disease or Alzheimer's disease), autonomic function
 CC disorders (e.g. hypertension, depression or schizophrenia), or learning
 CC or memory disorders (e.g. amnesia or age-related disorders),
 CC cardiovascular disorders (e.g. ischaemia reperfusion injury or
 CC restenosis), hormonal disorders (e.g. hypothyroidism or hyperthyroidism),
 CC or immune disorders (e.g. chronic mucocutaneous candidiasis)
 XX
 SQ Sequence 1095 AA;
 Query Match 82.8%; Score 24; DB 5; Length 1095;
 Best Local Similarity 71.4%; Pred. No. 2.8e+03;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 VMXVAEF 7
 :| |||||
 Db 1016 LMVVAEF 1022
 RESULT 74
 ABP52170
 ID ABP52170 standard; protein; 1095 AA.
 XX
 AC ABP52170;
 XX
 DT 14-OCT-2002 (first entry)
 XX
 DE Mouse potential phospholipid transporting ATPase IIV SEQ ID NO:36.
 XX
 DE Human; transporter protein; anorectic; antidiabetic; anti-Parkinsonian;
 KW neurotropic; neuroprotective; hypotensive; antidepressant; neuroleptic;
 KW cardiovascular; immunosuppressive; gene therapy; sugar homeostasis;
 KW obesity; diabetes; anorexia; central nervous system disorder; depression;
 KW CNS disorder; Parkinson's disease; Alzheimer's disease; hypertension;
 KW autonomic function disorder; schizophrenia; learning disorder; amnesia;
 KW memory disorder; age-related disorder; cardiovascular disorder;
 KW ischaemia reperfusion injury; restenosis; hormonal disorder;
 KW hypothyroidism; hyperthyroidism; immune disorder;
 KW chronic mucocutaneous candidiasis.
 XX
 OS Mus musculus.

XX Key Location/Qualifiers
 FH Misc-difference 801 /note= "unspecified"
 FT Misc-difference 1005 /note= "unspecified"
 FT
 XX WO200255701-A2.
 XX
 XX 18-JUL-2002.
 XX
 XX 17-DEC-2001; 2001WO-US049060.
 XX
 XX 15-DEC-2000; 2000US-0256240P.
 PR 18-DEC-2000; 2000US-0256588P.
 PR 21-DEC-2000; 2000US-0258028P.
 XX
 PA (MILL-) MILLENNIUM PHARM INC.
 XX
 PI Curtis RAJ;
 XX
 XX WPI; 2002-590672/63.
 DR
 XX New 67076, 67102, 44181, 67084FL and 67084alt nucleic acids and proteins,
 FT useful in e.g. treating disorders characterized by insufficient or
 PT excessive production of the polypeptides, e.g. cardiovascular or immune
 PT disorders.
 PT
 XX Disclosure; Fig 35; 364pp; English.
 PS
 XX The present invention describes human proteins which are members of the
 CC transporter family. The transporter family proteins (I) of the present
 CC invention have anorectic, antidiabetic, anti-Parkinsonian, nootropic,
 CC neuroprotective, hypotensive, antidepressant, neuroleptic, cardiovascular
 CC and immunosuppressive activities, and can be used in gene therapy. The
 CC human transporter proteins from the present invention are designated
 CC 8099, 46455, 54414, 53763, 67076, 67102, 44181, 67084FL and 67084alt. The
 CC polynucleotides encoding (I) can be used in screening assays (e.g.
 CC chromosome mapping, tissue typing, or in forensic biology), predictive
 CC medicine (e.g. diagnostic assays, prognostic assays, monitoring clinical
 CC trials, or pharmacogenetics), as surrogate markers, and in methods of
 CC treatment (e.g. therapeutic or prophylactic). (I) are useful for treating
 CC disorders characterised by insufficient or excessive production of 8099,
 CC 46455, 54414, 53763, 67076, 67102, 44181, 67084FL or 67084alt substrates
 CC or production of transport 8099, 46455, 54414, 53763, 67076, 67102,
 CC 44181, 67084FL or 67084alt inhibitors. Disorders associated with 8099 and
 CC 46455 include disorders associated with sugar homeostasis such as
 CC obesity, diabetes or anorexia. Disorders associated with 54414, 53763,
 CC 67076, 67102, 44181, 67084FL or 67084alt are e.g. CNS disorders (e.g.
 CC Parkinson's disease or Alzheimer's disease), autonomic function disorders
 CC (e.g. hypertension, depression or schizophrenia), or learning or memory
 CC disorders (e.g. amnesia or age-related disorders), cardiovascular
 CC disorders (e.g. ischaemia reperfusion injury or restenosis), hormonal
 CC disorders (e.g. hypothyroidism or hyperthyroidism), or immune disorders
 CC (e.g. chronic mucocutaneous candidiasis). The present sequence represents
 CC a protein given in comparison with a human transporter protein from the
 CC present invention
 XX
 XX Sequence 1095 AA;
 SQ
 Query Match 82.8%; Score 24; DB 5; Length 1095;
 Best Local Similarity 71.4%; Pred. No. 2.8e+03;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 VMKVAEF 7
 Db 1016 LMVVAEF 1022
 RESULT 75
 AAE32082
 ID AAE32082 standard; protein; 1095 AA.
 XX

AC AAE32082;
 XX 24-MAR-2003 (first entry)
 XX
 XX Human TRICH-16 protein.
 DE
 XX Human; transporter and ion channel; TRICH; atherosclerosis; cancer;
 KW gene therapy.
 XX
 XX Homo sapiens.
 OS
 XX
 FH Key Location/Qualifiers
 FT Domain 94..122
 FT /note= "Transmembrane domain"
 FT Domain 127..153
 FT /note= "Transmembrane domain"
 FT Domain 171..199
 FT /note= "E1-E2 ATPase domain"
 FT Domain 277..305
 FT /note= "E1-E2 ATPase domain"
 FT Domain 280..308
 FT /note= "Transmembrane domain"
 FT Domain 326..346
 FT /note= "Transmembrane domain"
 FT Domain 351..371
 FT /note= "Transmembrane domain"
 FT Domain 880..900
 FT /note= "Transmembrane domain"
 FT Domain 907..927
 FT /note= "Transmembrane domain"
 FT Domain 958..978
 FT /note= "Transmembrane domain"
 FT Domain 986..1006
 FT /note= "Transmembrane domain"
 FT Domain 1015..1035
 FT /note= "Transmembrane domain"
 FT Domain 1041..1061
 FT /note= "Transmembrane domain"
 XX WO200283712-A2.
 XX
 XX 24-OCT-2002.
 XX
 XX 12-APR-2002; 2002WO-US011760.
 XX
 XX 12-APR-2001; 2001US-0283440P.
 PR 20-APR-2001; 2001US-0285592P.
 PR 27-APR-2001; 2001US-0287263P.
 PR 04-MAY-2001; 2001US-0288666P.
 PR 18-MAY-2001; 2001US-0292042P.
 PR 25-MAY-2001; 2001US-0293724P.
 PR 22-JAN-2002; 2002US-0351107P.
 XX (INCY-) INCYTE GENOMICS INC.
 PA
 XX Baughn MR, Elliott VS, Hafalia AJA, Yang J, Wallia NK, Rankumar J;
 PI Forsythe IJ, Lu Y, Tang YT, Yue H, Raumann BE, Lal PG, Azimzai Y;
 PI Lu DAM, Gandhi AR, Thornton M, Nguyen DB, Arvizu CS, Emerling BM,
 PI Swarnakar A, Yao MG, Ding L, He A, Griffin JA, Sanjanwala MM;
 PI Gietzen KJ, Lee EA, Xu Y, Au-Young JK, Das D, Lee SY, Chang H;
 XX WPI; 2003-092996/08.
 DR N-PSDB; AAD49514.
 XX
 XX New human functional transporters and ion channels (TRICH) polypeptides,
 FT useful for preparing a composition for diagnosing or treating a disease
 PT associated with decreased expression or overexpression of TRICH e.g.
 PT cancer.
 XX
 XX Claim 1; Page 178-181; 204pp; English.
 PS
 XX The invention relates to human transporters and ion channels (TRICH)
 CC polypeptides and nucleic acid molecules encoding such polypeptides. TRICH

CC proteins are useful for preparing compositions for diagnosing or treating
 CC diseases or conditions associated with decreased expression or
 CC overexpression of functional TRICH e.g. atherosclerosis or cancer. The
 CC invention is useful in gene therapy. The present sequence is human TRICH
 CC protein
 XX
 SQ Sequence 1095 AA;
 Query Match 82.8%; Score 24; DB 6; Length 1095;
 Best Local Similarity 71.4%; Pred. No. 2.8e+03;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 VMXVAEF 7
 :| ||||
 Db 1016 LMVVAEF 1022
 RESULT 76
 ADD37509
 ID ADD37509 standard; protein; 1095 AA.
 XX
 AC ADD37509;
 XX
 DT 15-JAN-2004 (first entry)
 XX
 DE Mouse phospholipid transporting ATPase, AT2B.
 XX
 KW Transporter; cytostatic; anorectic; antidiabetic; anticonvulsant;
 KW gene therapy; PGC-1 associated disorder; liver tumour; obesity; epilepsy;
 KW diabetes.
 XX
 OS Mus sp.
 XX
 PN US2003143675-A1.
 XX
 PD 31-JUL-2003.
 XX
 PF 22-MAY-2002; 2002US-00154419.
 XX
 PR 12-MAY-2000; 2000US-0204211P.
 PR 29-JUN-2000; 2000US-0215376P.
 PR 31-JUL-2000; 2000US-0221769P.
 PR 19-SEP-2000; 2000US-0233790P.
 PR 25-SEP-2000; 2000US-0235107P.
 PR 05-OCT-2000; 2000US-0238336P.
 PR 14-NOV-2000; 2000US-0248364P.
 PR 15-NOV-2000; 2000US-0248878P.
 PR 15-DEC-2000; 2000US-0256240P.
 PR 18-DEC-2000; 2000US-0256588P.
 PR 21-DEC-2000; 2000US-0258028P.
 PR 22-JAN-2001; 2001US-0263169P.
 PR 14-MAY-2001; 2001US-0263169P.
 PR 29-JUN-2001; 2001US-00858194.
 PR 31-JUL-2001; 2001US-0085811.
 PR 19-SEP-2001; 2001US-00957664.
 PR 25-SEP-2001; 2001US-00964295.
 PR 05-OCT-2001; 2001US-00972724.
 PR 14-NOV-2001; 2001US-00002769.
 PR 17-DEC-2001; 2001US-00024623.
 PR 22-JAN-2002; 2002US-00055025.
 XX
 PA (MILL-) MILLENNIUM PHARM INC.
 XX
 PI Curtis RAJ, Gluckmann MA, Meyers RE;
 XX
 DR WPI; 2003-851783/79.
 XX
 XX New isolated nucleic acid, useful for preparing a composition for
 PT treating PGC-1 associated disorders e.g. liver tumors, obesity, epilepsy
 PT or diabetes.
 XX
 PS Example; SEQ ID NO 86; 663pp; English.
 XX

CC The invention relates to an isolated nucleic acid comprising a cDNA
 CC encoding a human transporter protein, or its complement, a sequence that
 CC is 60 % identical to the cDNA, a fragment comprising at least 30
 CC nucleotides of the cDNA, or a sequence encoding a fragment of the
 CC polypeptide comprising at least 10 contiguous amino acid residues of the
 CC cDNA. Also included are a vector comprising the novel nucleic acid
 CC molecule, producing the polypeptide, the isolated transporter
 CC polypeptide, an isolated antibody that specifically binds to the
 CC polypeptide, detecting the presence of the polypeptide or nucleic acid in
 CC a sample, a kit, identifying a compound that binds to, or that modulates
 CC the activity of, the polypeptide, and modulating the activity of the
 CC polypeptide. The nucleic acid is useful for preparing a composition for
 CC treating PGC-1 (not defined) associated disorders e.g. liver tumours,
 CC obesity, epilepsy or diabetes. The present sequence represents a species
 CC homologue of a novel human transporter protein of the invention.
 XX
 SQ Sequence 1095 AA;
 Query Match 82.8%; Score 24; DB 7; Length 1095;
 Best Local Similarity 71.4%; Pred. No. 2.8e+03;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 VMXVAEF 7
 :| ||||
 Db 1016 LMVVAEF 1022
 RESULT 77
 ADD37499
 ID ADD37499 standard; protein; 1095 AA.
 XX
 AC ADD37499;
 XX
 DT 15-JAN-2004 (first entry)
 XX
 DE Human transporter 67084alt.
 XX
 KW Human; transporter; cytostatic; anorectic; antidiabetic; anticonvulsant;
 KW gene therapy; PGC-1 associated disorder; liver tumour; obesity; epilepsy;
 KW diabetes.
 XX
 OS Homo sapiens.
 XX
 PN US2003143675-A1.
 XX
 PD 31-JUL-2003.
 XX
 PF 22-MAY-2002; 2002US-00154419.
 XX
 PR 12-MAY-2000; 2000US-0204211P.
 PR 29-JUN-2000; 2000US-0215376P.
 PR 31-JUL-2000; 2000US-0221769P.
 PR 19-SEP-2000; 2000US-0233790P.
 PR 25-SEP-2000; 2000US-0235107P.
 PR 05-OCT-2000; 2000US-0238336P.
 PR 14-NOV-2000; 2000US-0248364P.
 PR 15-NOV-2000; 2000US-0248878P.
 PR 15-DEC-2000; 2000US-0256240P.
 PR 18-DEC-2000; 2000US-0256588P.
 PR 21-DEC-2000; 2000US-0258028P.
 PR 22-JAN-2001; 2001US-0263169P.
 PR 14-MAY-2001; 2001US-00858194.
 PR 29-JUN-2001; 2001US-0085811.
 PR 31-JUL-2001; 2001US-00919781.
 PR 19-SEP-2001; 2001US-00957664.
 PR 25-SEP-2001; 2001US-00964295.
 PR 05-OCT-2001; 2001US-00972724.
 PR 14-NOV-2001; 2001US-00002769.
 PR 17-DEC-2001; 2001US-00024623.
 PR 22-JAN-2002; 2002US-00055025.
 XX
 PA (MILL-) MILLENNIUM PHARM INC.
 XX


```

PI Curtis RAJ, Gluckemann MA, Meyers RE;
XX WPI; 2003-851783/79.
DR N-PSDB; ADD37498, ADD37500.
XX
PT New isolated nucleic acid, useful for preparing a composition for
PT treating PGC-1 associated disorders e.g. liver tumors, obesity, epilepsy
PT or diabetes.
XX
PS Claim 11; SEQ ID NO 76; 663pp; English.
XX
CC The invention relates to an isolated nucleic acid comprising a cDNA
CC encoding a human transporter protein, or its complement, a sequence that
CC is 60 % identical to the cDNA, a fragment comprising at least 30
CC nucleotides of the cDNA, or a sequence encoding a fragment of the
CC polypeptide comprising at least 10 contiguous amino acid residues of the
CC cDNA. Also included are a vector comprising the novel nucleic acid
CC molecule, producing the polypeptide, the isolated transporter
CC polypeptide, an isolated antibody that specifically binds to the
CC polypeptide, detecting the presence of the polypeptide or nucleic acid in
CC a sample, a kit, identifying a compound that binds to, or that modulates
CC the activity of, the polypeptide, and modulating the activity of the
CC polypeptide. The nucleic acid is useful for preparing a composition for
CC treating PGC-1 (not defined) associated disorders e.g. liver tumors,
CC obesity, epilepsy or diabetes. The present sequence represents a novel
CC human transporter protein.
XX
SQ Sequence 1095 AA;

Query Match      82.8%; Score 24; DB 7; Length 1095;
Best Local Similarity 71.4%; Pred. No. 2.8e+03;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
Db 1016 LMVVAEF 1022

RESULT 78
AD127978
ID AD127978 standard; protein; 1095 AA.
XX
AC AD127978;
XX
DT 06-MAY-2004 (first entry)
XX
DE Human 67084alt protein.
XX
KW Human; 67084alt; ion channel family; ICF; cancer; leukaemia;
KW Alzheimer's disease; Parkinson's disease; multiple sclerosis; epilepsy;
KW hepatic disorder; cardiovascular disorder; cytosstatic; neuroprotective;
KW neurotropic; antiparkinsonian; hepatotropic; cardiovascular.
XX
OS Homo sapiens.
XX
PN US2003165891-A1.
XX
PD 04-SEP-2003.
XX
PF 15-MAY-2002; 2002US-00146733.
XX
PR 29-FEB-2000; 2000US-00515520.
PR 29-FEB-2000; 2000US-0185938P.
PR 03-MAR-2000; 2000US-00518866.
PR 07-APR-2000; 2000US-0195734P.
PR 11-APR-2000; 2000US-0195993P.
PR 26-APR-2000; 2000US-0199799P.
PR 19-SEP-2000; 2000US-0233537P.
PR 25-SEP-2000; 2000US-0235018P.
PR 15-SEP-2000; 2000US-0235059P.
PR 25-SEP-2000; 2000US-0256240P.
PR 18-DEC-2000; 2000US-0256588P.
PR 21-DEC-2000; 2000US-0258028P.
XX

```

```

PR 28-FEB-2001; 2001US-00796720.
PR 06-APR-2001; 2001US-00828035.
PR 11-APR-2001; 2001US-00833081.
PR 25-APR-2001; 2001US-00843128.
PR 19-SEP-2001; 2001US-00957683.
PR 25-SEP-2001; 2001US-00964252.
PR 15-SEP-2001; 2001US-00964256.
PR 17-DEC-2001; 2001US-00024623.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Curtis RAJ, Gluckemann MA, Silos-Santiago I;
XX WPI; 2004-069000/07.
DR N-PSDB; AD127977, AD127979.
XX
XX TWIK-6, TWIK-7, IC23927, TWIK-8, IC47611, IC47615, HNMDA-1, TWIK-9,
PT alpha2delta-4, 54414, or 53763 nucleic acids and proteins, useful for
PT preventing, diagnosing and treating e.g. cancers, Alzheimer's disease and
PT cardiovascular disorders.
XX
PS Example 1; SEQ ID NO 71; 638pp; English.
XX
CC The invention relates to TWIK-6, TWIK-7, IC23927, TWIK-8, IC47611,
CC IC47615, HNMDA-1, TWIK-9, alpha-2delta-4, 54414 and 53763 ion channel
CC family (ICF) nucleic acids and proteins. The TWIK-6, TWIK-7, IC23927,
CC TWIK-8, IC47611, IC47615, HNMDA-1, TWIK-9, alpha-2delta-4, 54414 and
CC 53763 ICF nucleic acids and proteins may be used for preventing,
CC diagnosing and treating ICF-related diseases. The sequences may be used
CC to treat disorders associated with decreased expression by rectifying
CC mutations or deletions in a patient's genome that affect the activity of
CC ICF proteins by expressing inactive proteins or to supplement the
CC patients own production of ICF proteins. The proteins may also be used as
CC antigens in the production of antibodies against ICF proteins and in
CC assays to identify modulators of ICF protein expression and activity. The
CC anti-ICF protein antibodies, agonists and antagonists may be used to
CC regulate ICF protein expression and activity. The antibodies may also be
CC used as diagnostic agents for detecting the presence of ICF proteins in
CC samples (e.g. by immunoassay). The nucleic acids and proteins may be used
CC to prevent, diagnose and treat a wide variety of disorders, e.g. cancers
CC and leukaemia, Alzheimer's disease, Parkinson's disease, multiple
CC sclerosis, epilepsy, hepatic disorders and cardiovascular disorders. This
CC sequence represents the human 67084alt protein of the invention. Note:
CC The sequence data for this patent is also available in electronic format
CC from USPTO at seqdata.uspto.gov/sequence.html.
XX
SQ Sequence 1095 AA;

Query Match      82.8%; Score 24; DB 8; Length 1095;
Best Local Similarity 71.4%; Pred. No. 2.8e+03;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
Db 1016 LMVVAEF 1022

RESULT 79
AD127988
ID AD127988 standard; protein; 1095 AA.
XX
AC AD127988;
XX
DT 06-MAY-2004 (first entry)
XX
DE Murine protein #4.
XX
KW Mouse; ion channel family; ICF; cancer; leukaemia; Alzheimer's disease;
KW Parkinson's disease; multiple sclerosis; epilepsy; hepatic disorder;
KW cardiovascular disorder; cytosstatic; neuroprotective; neurotropic;
KW antiparkinsonian; hepatotropic; cardiovascular.
XX
OS Mus musculus.

```


XX	OS	Rattus sp.	XX	PS	Example 2; Col 51-52; 46pp; English.
XX	PN	US5811516-A.	XX	CC	This represents a receptor protein tyrosine kinase (PTK) subtype tyro-9.
XX	XX		XX	CC	The invention provides sequences AAV65308 to AAV65313, AAV65315, and
XX	PD	22-SEP-1998.	XX	CC	AAV65317 to AAV65319 that encode proteins having a tyrosine kinase domain
XX	XX		XX	CC	and a tissue expression pattern of a receptor PTK subtype selected from
XX	PF	02-JUN-1995; 95US-00456647.	XX	CC	tyro-1, tyro-2, tyro-3, tyro-4, tyro-5, tyro-6, tyro-8, tyro-10, tyro-11,
XX	XX		XX	CC	and tyro-12, respectively. The polynucleotides are useful for the
XX	PR	15-MAY-1992; 92US-00884486.	XX	CC	detection of tyrosine kinase domain sequences and detection of tissue
XX	PR	02-MAY-1994; 94US-00237401.	XX	CC	expression patterns of PTK subtypes. The cDNAs can also be injected into
XX	XX	(SALK) SALK INST BIOLOGICAL STUDIES.	XX	CC	oocytes, the protein expressed, and expression products screened for
XX	XX		XX	CC	using antibodies against tyrosine kinase epitopes. These subtypes
XX	PI	Lai CHC, Lemke GE;	XX	CC	sequences can be used for the design of oligonucleotides, for use in
XX	XX		XX	CC	amplification reactions to isolate other subtype sequences. These
XX	DR	WPI; 1998-530939/45.	XX	CC	detection protocols are used in the diagnosis of diseases associated with
XX	DR	N-PSDB; AAV5894.	XX	CC	(receptor) PTKs. Recombinant vectors expressing the subtypes can be used
XX	XX		XX	CC	to treat related diseases e.g. tumours, by introduction of the vectors
XX	PT	Receptor protein tyrosine kinase polypeptide, tyro-3 - preferably	XX	CC	into skin transplants, then grafting these into the connective tissue of
XX	PT	expressed in brain tissue.	XX	CC	the dermis, thus specifically targeting tumours as the proteins are
XX	XX		XX	CC	released from the matrix
XX	PS	Example 2; Col 51-52; 46pp; English.	XX	SQ	Sequence 54 AA;
XX	CC				
XX	CC	This represents a novel receptor protein tyrosine kinase (PTK)			
XX	CC	polypeptide subtype tyro-9. The invention provides polynucleotide			
XX	CC	sequences encoding novel PTK polypeptide subtypes tyro-1 to tyro-13. The			
XX	CC	PTK subtypes are found expressed predominantly in the brain tissue			
XX	XX				
XX	SQ	Sequence 54 AA;			
Query Match 79.3%; Score 23; DB 2; Length 54;			Query Match 79.3%; Score 23; DB 2; Length 54;		
Best Local Similarity 57.1%; Pred. No. 2e+02;			Best Local Similarity 57.1%; Pred. No. 2e+02;		
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;			Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;		
QY	1	VMXVAEF 7	QY	1	VMXVAEF 7
		:			:
Db	9	VMKIADF 15	Db	9	VMKIADF 15
RESULT 84			RESULT 85		
AAW81408			ABU83035		
ID	AAW81408	standard; protein; 54 AA.	ID	ABU83035	standard; protein; 54 AA.
XX	XX		XX	XX	
AC	AAW81408;		AC	ABU83035;	
XX	XX		XX	XX	
DT	22-JAN-1999	(first entry)	DT	27-JUN-2003	(first entry)
XX	XX		XX	XX	
DE	Receptor protein tyrosine kinase (PTK) subtype tyro-9.		DE	Mouse receptor protein tyrosine kinase bek.	
XX	XX		XX	XX	
KW	PTK; receptor; protein tyrosine kinase; recombinant; grafting; diagnosis;		XX	XX	
KW	tumour; skin transplant; connective tissue; tyro-9.		KW	Protein tyrosine kinase; PTK receptor; PTK receptor modulation assay;	
XX	XX		XX	mouse.	
OS	Rattus sp.		XX	Mus sp.	
XX	XX		XX	XX	
PN	US5837448-A.		PN	US2003013848-A1.	
XX	XX		XX	XX	
PD	17-NOV-1998.		PD	16-JAN-2003.	
XX	XX		XX	XX	
XX	02-MAY-1994; 94US-00237401.		XX	XX	
XX	XX		XX	XX	
XX	15-MAY-1992; 92US-00884486.		XX	XX	
XX	XX		XX	XX	
PA	(SALK) SALK INST BIOLOGICAL STUDIES.		XX	XX	
XX	XX		XX	XX	
PI	Lai CHC, Lemke GE;		XX	XX	
XX	XX		XX	XX	
XX	WPI; 1999-023436/02.		XX	XX	
XX	N-PSDB; AAV65316.		XX	XX	
XX			XX	XX	
PT	Nucleic acids encoding protein tyrosine kinase subtypes - for		XX	XX	
PT	identification of new sub:types and treatment of diseases associated with		XX	XX	
PT	the kinase.		XX	XX	

CC protein(s) comprises a tyrosine kinase domain and a tissue expression
 CC pattern having at least one receptor protein-tyrosine kinase subtype
 CC selected from tyro-1, tyro-2, tyro-4, tyro-5, tyro-6, tyro-7, tyro-8,
 CC tyro-10, tyro-11 and tyro-12. The protein is useful in assays to screen
 CC various compositions which modulate these protein-tyrosine kinase
 CC receptors. The presents sequence represents the amino acid sequence of a
 CC receptor protein tyrosine kinase domain
 XX
 SQ Sequence 54 AA;

Query Match 79.3%; Score 23; DB 6; Length 54;
 Best Local Similarity 57.1%; Pred. No. 2e+02;
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
 || :||
 Db 9 VKIADF 15

RESULT 86

ABU83019
 ID ABU83019 standard; protein; 54 AA.

XX AC ABU83019;

XX DT 27-JUN-2003 (first entry)

DE Rat receptor protein tyrosine kinase tyro-9.

XX Protein tyrosine kinase; PTK receptor; PTK receptor modulation assay;
 KW rat.

XX OS Rattus sp.

XX PN US2003013848-A1.

XX PD 16-JAN-2003.

XX PF 22-SEP-1998; 98US-00158722.

XX PR 15-MAY-1992; 92US-00884486.

XX PR 02-MAY-1994; 94US-00237401.

XX PR 02-JUN-1995; 95US-00456647.

XX PA (LEMK/) LEMKE G E.

XX PA (LAIC/) LAI C H C.

XX PI Lemke GE, Lai CHC;

XX DR WPI; 2003-401606/38.

XX DR N-PSDB; ACA69713.

XX PT Novel pure protein-tyrosine kinase receptor subtypes useful in assays to
 PT screen various compositions which modulate these receptors.

XX PS Example 2; Page 12; 49pp; English.

XX The invention relates to a substantially pure protein(s) or their
 CC fragments, comprising a tyrosine kinase domain and a tissue expression
 CC pattern having at least one receptor protein-tyrosine kinase subtype. The
 CC protein(s) comprises a tyrosine kinase domain and a tissue expression
 CC pattern having at least one receptor protein-tyrosine kinase subtype
 CC selected from tyro-1, tyro-2, tyro-4, tyro-5, tyro-6, tyro-7, tyro-8,
 CC tyro-10, tyro-11 and tyro-12. The protein is useful in assays to screen
 CC various compositions which modulate these protein-tyrosine kinase
 CC receptors. The presents sequence represents the amino acid sequence of a
 CC receptor protein tyrosine kinase tyro subtype
 XX

SQ Sequence 54 AA;

Query Match 79.3%; Score 23; DB 6; Length 54;
 Best Local Similarity 57.1%; Pred. No. 2e+02;
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
 || :||
 Db 9 VKIADF 15

RESULT 87

ABU83034
 ID ABU83034 standard; protein; 54 AA.

XX AC ABU83034;

XX DT 27-JUN-2003 (first entry)

XX Receptor protein tyrosine kinase bFGF-R.

XX Protein tyrosine kinase; PTK receptor; PTK receptor modulation assay.

XX OS Unidentified.

XX PN US2003013848-A1.

XX PD 16-JAN-2003.

XX PF 22-SEP-1998; 98US-00158722.

XX PR 15-MAY-1992; 92US-00884486.

XX PR 02-MAY-1994; 94US-00237401.

XX PR 02-JUN-1995; 95US-00456647.

XX PA (LEMK/) LEMKE G E.

XX PA (LAIC/) LAI C H C.

XX PI Lemke GE, Lai CHC;

XX DR WPI; 2003-401606/38.

XX PT Novel pure protein-tyrosine kinase receptor subtypes useful in assays to
 PT screen various compositions which modulate these receptors.

XX PS Example 2; Page 12; 49pp; English.

XX The invention relates to a substantially pure protein(s) or their
 CC fragments, comprising a tyrosine kinase domain and a tissue expression
 CC pattern having at least one receptor protein-tyrosine kinase subtype. The
 CC protein(s) comprises a tyrosine kinase domain and a tissue expression
 CC pattern having at least one receptor protein-tyrosine kinase subtype
 CC selected from tyro-1, tyro-2, tyro-4, tyro-5, tyro-6, tyro-7, tyro-8,
 CC tyro-10, tyro-11 and tyro-12. The protein is useful in assays to screen
 CC various compositions which modulate these protein-tyrosine kinase
 CC receptors. The presents sequence represents the amino acid sequence of a
 CC receptor protein tyrosine kinase domain
 XX

SQ Sequence 54 AA;

Query Match 79.3%; Score 23; DB 6; Length 54;
 Best Local Similarity 57.1%; Pred. No. 2e+02;
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
 || :||
 Db 9 VKIADF 15

RESULT 88

AAU78556
 ID AAU78556 standard; protein; 56 AA.

XX AC AAU78556;

XX DT 18-JUN-2002 (first entry)

XX Residues 592-647 of human fibroblast growth factor receptor 1 (FGFR1).

XX Tie2; Tie2K; human; crystal structure; vascular dysmorphogenesis;
KW protein co-ordinate data; receptor tyrosine kinase; FGFR1;
KW vascular development; fibroblast growth factor receptor.
XX
OS Homo sapiens.
XX
PN WO200220734-A2.
XX
PD 14-MAR-2002.
XX
XX 06-SEP-2001; 2001WO-US027486.
PF
XX 08-SEP-2000; 2000US-0231398P.
PR
XX (GLAX) GLAXO GROUP LTD.
PA
XX Davis RG, Ellis BP, Hassell AM, Holmes WD, Shewchuk LM;
PI WPI; 2002-292263/33.
XX
XX Composition for the design or screening of a cytoplasmic Tie2 receptor
PT tyrosine kinase domain modulator, comprises a monoclinic or orthorhombic
PT crystalline form of a cytoplasmic Tie2 receptor tyrosine kinase domain
PT polypeptide.
XX
XX Disclosure; Page 579; 581pp; English.
PS
XX This invention relates to the 3 dimensional crystal structure of the
CC human Tie2 receptor tyrosine kinase domain polypeptide in a monoclinic,
CC or orthorhombic crystalline form. Tie2 is an endothelial-specific
CC receptor tyrosine kinase thought to be involved in vascular development.
CC Naturally occurring R849W and Y897S mutations in the Tie2 protein have
CC been identified in human s and have been shown to segregate with the
CC autosomal dominant condition vascular dysmorphogenesis. The invention
CC also comprises methods for designing modulators of the biological
CC activity of the cytoplasmic Tie2 receptor tyrosine kinase domain. The
CC Tie2 receptor tyrosine kinase domain is used to design or screen for a
CC modulator of the kinase by rational drug design, using computer models.
CC The present sequence represents residues 592-647 of human fibroblast
CC growth factor receptor 1 (FGFR1). The catalytic domain of FGFR1 closely
CC resembles that of Tie 2 and they share 45 percent sequence identity
XX
SQ Sequence 56 AA;

Query Match 79.3%; Score 23; DB 5; Length 56;
Best Local Similarity 57.1%; Pred. No. 2.1e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VMXVAEF 7
Db 45 VMKIADF 51

RESULT 89
ABF31047
ID ABP31047 standard; protein; 68 AA.
XX
XX ABP31047;
AC
XX
DT 08-JUL-2002 (first entry)
XX
DE Human ORF20 protein, SEQ ID NO:40.
XX
KW Human; ORF; open reading frame; ORF; drug screening; diagnosis;
KW disease monitoring; cytokine; cell proliferation; cell differentiation;
KW immune modulation; haematopoiesis regulation; tissue growth;
KW angiogenesis; activin; inhibin; chemotactic; chemokinetic; haemostatic;
KW thrombolytic; tumour inhibition; bodily characteristic; fertility;
KW behaviour; cancer; proliferative disorder; neurological disorder;
KW cardiovascular disease; immune system disorder; organ transplantation;
KW tissue growth disorder; tissue regeneration disorder; diabetes mellitus;
KW hypothyroidism; cholesterol ester storage disease; infection; vulnery;

KW vasotropic; antipsoriatic; antidiabetic; cytostatic; nootropic;
KW neuroprotective; antiatherosclerotic; anticoagulant; thrombolytic;
KW cardiant; hypotensive; antithyroid; antiinflammatory; immunomodulator;
KW dermatological; analgesic; virucide; antibacterial; fungicide.
XX
OS Homo sapiens.
XX
XX WO200190366-A2.
PN
XX
PD 29-NOV-2001.
XX
XX 24-MAY-2001; 2001WO-US017076.
PF
XX 24-MAY-2000; 2000US-0206690P.
PR
XX (CURA-) CURAGEN CORP.
PA
XX Leach MD, Shinkets RA;
PI WPI; 2002-106200/14.
XX N-PSDB; ABN75073.
DR
XX Novel human polypeptides and polynucleotides useful for diagnosing,
PT preventing and treating cardiovascular disease, neurodegenerative,
PT hyperproliferative disorders and disorders related to organ
PT transplantation.
XX
XX Claim 10; Page 280; 2508pp; English.
PS
XX Sequences ABP31028-ABP35561 represent 4534 novel human proteins
CC designated ORF (open reading frame) 1-4534, and sequences ABN75054-
CC ABN79587 represent cDNAs encoding them. The invention also encompasses
CC polypeptides at least 80% identical to the ORF1-ORF4534 (collectively
CC referred to as ORFX) proteins, polynucleotides at least 85% identical to
CC the ORFX nucleic acid sequences, vectors and host cells comprising ORFX
CC polynucleotides, the recombinant production of ORFX proteins, antibodies
CC specific for ORFX proteins, methods of detecting ORFX polynucleotides and
CC polypeptides, methods of screening for modulators of ORFX expression or
CC activity, and methods of screening individuals for a predisposition to an
CC ORFX-associated disorder. The ORFX proteins of the invention have a wide
CC range of biological activities, such as cytokine, cell proliferation,
CC cell differentiation, immune modulation, haematopoiesis regulation,
CC tissue growth, angiogenesis, activin or inhibin activity, chemotactic/
CC chemokinetic activity, haemostatic activity, thrombolytic activity,
CC receptor/ligand, antiinflammatory activity, tumour inhibition activity,
CC and antifinfective activity, and may also be involved in the determination
CC of bodily characteristics, fertility and behaviour. ORFX proteins,
CC nucleic acids and antibodies may be used in the treatment of cancers,
CC other proliferative disorders such as psoriasis and benign tumours,
CC neurological disorders such as epilepsy and Alzheimer's disease,
CC cardiovascular diseases, immune system disorders, disorders related to
CC organ transplantation, disorders of tissue growth and regeneration,
CC diseases such as diabetes mellitus, hypothyroidism, and cholesterol ester
CC storage disease, and infectious diseases caused by viral, bacterial,
CC fungal and other pathogens. ORFX nucleic acids may also be used as a
CC source of primers and probes, in the detection of ORFX genomic sequences
CC or transcripts, in the identification and cloning of homologous
CC sequences, in genetic diagnosis, and in forensic biology. The ORFX
CC nucleic acids may additionally be used to produce transgenic animals
CC which may be useful for studying the function and/or activity of ORFX
CC protein, and in drug screening. The ORFX proteins may also be used as
CC immunogens to generate specific antibodies, which are useful in the
CC diagnosis, treatment and monitoring of ORFX-associated diseases
XX
SQ Sequence 68 AA;

Query Match 79.3%; Score 23; DB 5; Length 68;
Best Local Similarity 83.3%; Pred. No. 2.6e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 MXVAEF 7
Db 1 MLVAEF 6

```
RESULT 90
AAU41413
ID AAU41413 standard; protein; 84 AA.
XX AC AAU41413;
XX DT 13-FEB-2002 (first entry)
XX DE Propionibacterium acnes immunogenic protein #2309.
XX KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
XX KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
XX KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
XX KW dermatological; osteopathic; neuroprotectant.
XX OS Propionibacterium acnes.
XX PN WO200181581-A2.
XX PD 01-NOV-2001.
XX PF 20-APR-2001; 2001WO-US012865.
XX PR 21-APR-2000; 2000US-0199047P.
XX PR 02-JUN-2000; 2000US-0208841P.
XX PR 07-JUL-2000; 2000US-0216747P.
XX PA (CORI-) CORIXA CORP.
XX PI Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
XX PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX DR WPI; 2001-616774/71.
XX DR N-PSDB; AAS9515.
XX PT Propionibacterium acnes polypeptides and nucleic acids useful for
XX PT vaccinating against and diagnosing infections, especially useful for
XX PT treating acne vulgaris.
XX PS Example 1; SEQ ID NO 2608; 1069pp; English.
XX CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
XX CC polypeptides. The proteins and their associated DNA sequences are used in
XX CC the treatment, prevention and diagnosis of medical conditions caused by
XX CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
XX CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
XX CC P. acnes is also involved in infections of bone, joints and the central
XX CC nervous system, however it is particularly involved in the inflammatory
XX CC lesions associated with acne vulgaris. A method for detecting the
XX CC presence or absence of P. acnes in a patient comprises contacting a
XX CC sample with a binding agent that binds to the proteins of the invention
XX CC and determining the amount of bound protein in the sample. The
XX CC polypeptides may be used as antigens in the production of antibodies
XX CC specific for P. acnes proteins. These antibodies can be used to
XX CC downregulate expression and activity of P. acnes polypeptides and
XX CC therefore treat P. acnes infections. The antibodies may also be used as
XX CC diagnostic agents for determining P. acnes presence, for example, by
XX CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
XX CC this patent did not form part of the printed specification, but was
XX CC obtained in electronic format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 84 AA;
Query Match 79.3%; Score 23; DB 4; Length 84;
Best Local Similarity 57.1%; Pred. No. 3.2e+02;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
Qy 1 VMKVAEP 7
Db 72 VMAIGEP 78

RESULT 91
ABM37932
ID ABM37932 standard; protein; 84 AA.
XX AC ABM37932;
XX DT 20-OCT-2003 (first entry)
XX DE Propionibacterium acnes predicted ORF-encoded polypeptide #2608.
XX KW Acne vulgaris; antiseborrheic; dermatological; antibacterial;
XX KW immunostimulant; immune response; vaccine.
XX OS Propionibacterium acnes.
XX PN WO2003033515-A1.
XX PD 24-APR-2003.
XX PF 11-OCT-2002; 2002WO-US032727.
XX PR 15-OCT-2001; 2001US-00978825.
XX PA (CORI-) CORIXA CORP.
XX PI Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;
XX PI Zhang Y, Wang S, Jen S, Lodes MD, Benson DR, Jones R, Carter D;
XX PI Barth B, Valliave-Douglas J;
XX DR WPI; 2003-381789/36.
XX DR N-PSDB; ACF64444.
XX PT New Propionibacterium acnes polypeptides and polynucleotides encoding the
XX PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
XX PT or for stimulating an immune response specific for a P. acnes protein.
XX PS Example 1; SEQ ID NO 2608; 1481pp; English.
XX CC The invention relates to an isolated polynucleotide (ACF64435-ACF64733)
XX CC encoding a Propionibacterium acnes protein. The invention also relates to
XX CC polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to
XX CC immunogenic fragments of P. acnes polypeptides. The invention
XX CC additionally encompasses expression vectors and host cells comprising a
XX CC polynucleotide of the invention, antibodies against polypeptides of the
XX CC invention; fusion proteins comprising a polypeptide of the invention; a
XX CC method for stimulating an immune response specific for a P. acnes
XX CC polypeptide and an isolated T cell population comprising T cells prepared
XX CC via this method; a vaccine composition (comprising P. acnes polypeptides,
XX CC polynucleotides, antibodies, fusion proteins, T cell populations, or
XX CC antigen-presenting cells that express the polypeptide); a method and kit
XX CC for detecting or determining the presence or absence of P. acnes in a
XX CC patient; and a method for inhibiting the development of P. acnes in a
XX CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion
XX CC proteins, T cell populations or antigen-presenting cells that express the
XX CC polypeptides are useful for diagnosing, preventing or treating acne
XX CC vulgaris, or for stimulating an immune response specific for a P. acnes
XX CC protein. The polynucleotides can also be used as probes or primers for
XX CC nucleic acid hybridisation. The vaccine composition is useful for the
XX CC stimulation of an immune response against P. acnes, or for treating acne,
XX CC and the kit is useful for performing a diagnostic assay. The present
XX CC sequence represents a polypeptide predicted to be encoded by an ORF (open
XX CC reading frame) contained within the P. acnes polynucleotides of the
XX CC invention. Note: The sequence data for this patent did not form part of
XX CC the printed specification, but was obtained in electronic format directly
XX CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 84 AA;
Query Match 79.3%; Score 23; DB 6; Length 84;
Best Local Similarity 57.1%; Pred. No. 3.2e+02;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
```

```

QY      1 VMXVAEF 7
Db      || : |||
       72 VMAIGEF 78

RESULT 92
AAB87752
ID AAB87752 standard; protein; 90 AA.
AC AAB87752;
XX
XX
DT 16-MAY-2001 (first entry)
XX
DE Human T2R22 amino acid sequence SEQ ID NO:40.
XX
XX Human; rat; mouse; T2R; taste receptor; G-protein coupled receptor;
KW taste transduction G-protein coupled receptor; identification; tongue;
KW taste sensory neuron; taste cell; taste modulator; food;
KW taste signalling pathway.
XX
XX Homo sapiens.
OS
XX WO200118050-A2.
XX
XX 15-MAR-2001.
XX
XX 08-SEP-2000; 2000WO-US024821.
XX
XX 10-SEP-1999; 99US-00393634.
XX
XX 22-FEB-2000; 2000US-00510332.
XX
XX (REGC ) UNIV CALIFORNIA.
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
XX Zukoer CS, Adler JE, Ryba N, Mueller K, Hoon M;
XX WPI; 2001-211396/21.
XX N-PSDB; AAF92520.
XX
XX Nucleic acids encoding the T2R family of G-protein coupled taste
PT receptors, useful for identifying taste modulators that can be used in
PT food and pharmaceutical industries to customize taste, for e.g. to
PT decrease the bitter taste of food.
XX
XX Claim 19; Page 177; 249pp; English.
XX
XX AAF92502 to AAF92572 represent nucleic acids which encode taste
CC transduction G-protein coupled receptors designated T2R proteins.
CC AAB87731 to AAB87824 represent T2R proteins, and AAB87825 to AAB87830
CC represent T2R family consensus sequences from the present invention. The
CC T2R proteins are taste modulators. The nucleic acids are useful as probes
CC for the identification of taste cells, as the nucleic acids are
CC specifically expressed in taste cells. They also serve as tools for the
CC generation of taste topographic maps that elucidate the relationship
CC between the taste cells of the tongue and taste sensory neurons leading
CC to taste centres in the brain. The taste modulators are useful for
CC pharmacological and genetic modulation of taste signalling pathways.
CC Modulatory compounds comprising T2R proteins can therefore be used in
CC food and pharmaceutical industries to customise taste, for e.g. to
CC decrease the bitter taste of food or drugs
XX
XX Sequence 90 AA;
SQ
Query Match 79.3%; Score 23; DB 4; Length 90;
Best Local Similarity 42.9%; Pred. No. 3.5e+02;
Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY      1 VMXVAEF 7
Db      || : |||
       11 ILAIAEF 17

```

```

RESULT 93
ADJ84523
ID ADJ84523 standard; protein; 90 AA.
XX
XX AC ADJ84523;
XX
XX 20-MAY-2004 (first entry)
XX
DE Human T2R G-protein coupled receptor seq id 76.
XX
XX taste transduction; G-protein coupled receptor; T2R; taste signaling;
KW CAMP level; cGMP level; IP3 level; Ca(2+) level; electrical activity;
KW human; G-protein coupled receptor; receptor.
XX
XX Homo sapiens.
OS
XX US2004038312-A1.
XX
XX 26-FEB-2004.
PD
XX 10-FEB-2003; 2003US-00364861.
XX
XX 10-SEP-1999; 99US-00393634.
XX
XX (ZUKER/) ZUKER C S.
XX (ADLER/) ADLER J E.
XX (HOON/) HOON M.
XX (RYBA/) RYBA N.
XX (MUELLER/) MUELLER K.
XX
XX Zukoer CS, Adler JE, Hoon M, Ryba N, Mueller K;
XX WPI; 2004-203221/19.
XX N-PSDB; ADJ84524.
XX
XX Novel isolated taste transduction G-protein coupled receptor e.g., T2R
PT useful for identifying compound that modulates taste signaling in taste
PT cells.
XX
XX Claim 22; SEQ ID NO 76; 121pp; English.
XX
XX The invention describes an isolated taste transduction G-protein coupled
CC receptor (I) e.g., T2R which is expressed in a taste cell, comprises
CC greater than 60% sequence identity to a fully defined sequence of 335
CC (S1), 333 (S2), 299 (S3), 310 (S4), 224 (S5), 77 (S6), 209 (S7), 266
CC (S8), 300 (S9), 300 (S10), 155 (S11), 173 (S12), 94 (S13), 115 (S14), 68
CC (S15), 126 (S16) and 180 (S17) amino acids as given in the specification.
CC (I) is useful for identifying a compound that modulates taste signaling
CC in taste cells which involves contacting the compound with (I) and
CC determining the functional effect of the compound (I). The functional
CC effect is determined by measuring changes in intracellular cAMP, cGMP,
CC IP3, or Ca (2+). The functional effect is a chemical or physical effect.
CC The functional effect is determined by measuring binding of the compound
CC to an extracellular domain of (I). The functional effect is determined by
CC measuring binding of radiolabeled GTP to (I). (I) is recombinant and is
CC from rat, mouse, or human. (I) is expressed in a cell or cell membrane.
CC The functional effect is measured by determining changes in the
CC electrical activity of cells (e.g., eukaryotic cell) expressing (I). (I)
CC is useful for identifying compound that modulates taste signaling in
CC taste cells which involves contacting a compound with (I). The
CC extracellular domain (II) of (I) is useful for identifying a compound
CC that modulates taste signaling in taste cells which involves contacting
CC the compound with the polynucleotide encoding (I) and determining the
CC functional effect of the compound upon the extracellular domain. This
CC sequence encodes a human T2R G-protein coupled receptor.
XX
XX Sequence 90 AA;
SQ
Query Match 79.3%; Score 23; DB 8; Length 90;
Best Local Similarity 42.9%; Pred. No. 3.5e+02;
Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY      1 VMXVAEF 7

```


CC (e.g. for immunotherapy of human immune deficiency virus infection or
 CC tumors) and in vaccines, including multivalent vaccines, against
 CC bacterial or viral infections, to produce diagnostic or therapeutic
 CC antibodies, for selective detection, purification and characterization of
 CC antibodies, and for preparation of protein libraries. (I) may also be
 CC used as biosensors and for diagnosis, e.g. of tumors. Genes that encode
 CC (I) are useful in DNA vaccines and for preparation of plant-based oral
 CC vaccines. (I) can contain many FR (same or different) at the surface of a
 CC spherical particle (LS comprises 60 subunits that assemble into an
 CC icosahedron). The large number of FR may increase sensitivity in
 CC immunoassays and the efficiency of immunotherapy agents
 CC
 XX Sequence 143 AA;
 SQ

Query Match 79.3%; Score 23; DB 3; Length 143;
 Best Local Similarity 83.3%; Pred. No. 5.7e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 2 MVVAEF 7
 Db 8 MVVAEF 13

RESULT 99
 AAU33403
 ID AAU33403 standard; protein; 150 AA.
 XX
 AC AAU33403;
 XX
 DT 14-FEB-2002 (first entry)
 DE
 DE Enterococcus faecalis cellular proliferation protein #39.
 XX
 KW Antisense; prokaryotic cellular proliferation protein; antibiotic;
 KW antibacterial; drug design.
 XX
 OS Enterococcus faecalis.
 XX
 PN WO200170955-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 21-MAR-2001; 2001WO-US009180.
 XX
 PR 21-MAR-2000; 2000US-0191078P.
 PR 23-MAY-2000; 2000US-0206848P.
 PR 26-MAY-2000; 2000US-0207727P.
 PR 23-OCT-2000; 2000US-0242578P.
 PR 27-NOV-2000; 2000US-0253625P.
 PR 22-DEC-2000; 2000US-0257931P.
 PR 16-FEB-2001; 2001US-0269308P.
 XX
 PA (ELIT-) ELITRA PHARM INC. .
 XX
 PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
 PI Yamamoto RT, Xu HH;
 XX
 DR WPI; 2001-611495/70.
 DR N-PSDB; AAS51262.
 XX
 PT New polynucleotides for the identification and development of
 PT antibiotics, comprise sequences of antisense nucleic acids.
 XX
 PS Example 3; SEQ ID NO 4899; 511pp; English.
 XX
 CC The invention relates to antisense inhibitors of genes essential to
 CC prokaryotic cellular proliferation, their use in identifying the genes,
 CC their use in the discovery of novel antibiotics, the essential genes
 CC themselves and the encoded proteins. The prokaryotes used are Escherichia
 CC coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
 CC Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
 CC useful for the identification of potential new targets for antibiotic
 CC development. The antisense nucleic acids can also be used to identify

CC proteins used in proliferation, to express these proteins, and to obtain
 CC antibodies capable of binding to the expressed proteins. The proteins can
 CC be used to screen compounds in rational drug discovery programmes. The
 CC antisense nucleic acid sequence is also useful to screen for homologous
 CC nucleic acids which are required for cell proliferation in a wide variety
 CC of organisms. The present sequence represents an essential prokaryotic
 CC cellular proliferation protein. Note: The sequence data for this patent
 CC did not form part of the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 150 AA;

Query Match 79.3%; Score 23; DB 4; Length 150;
 Best Local Similarity 57.1%; Pred. No. 6e+02;
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 1 VMXVAEF 7
 Db 40 LMDIAEF 46

RESULT 100
 ABB55092
 ID ABB55092 standard; protein; 162 AA.
 XX
 AC ABB55092;
 XX
 DT 29-AUG-2003 (revised)
 DT 16-MAY-2002 (first entry)
 XX
 DE Lactococcus lactis protein arok.
 XX
 KW Biosynthesis; biodegradation; lactic bacterium; yogurt; cheese.
 XX
 OS Lactococcus lactis; IL1403.
 XX
 PN FR2807446-A1.
 XX
 PD 12-OCT-2001.
 XX
 PF 11-APR-2000; 2000FR-00004630.
 XX
 PR 11-APR-2000; 2000FR-00004630.
 XX
 PA (INRG) INRA INST NAT RECH AGRONOMIQUE.
 XX
 PI Bolotine A, Sorokine A, Renault P, Ehrlich SD;
 XX WPI; 2002-043418/06.
 XX
 PT New nucleotide sequence useful in the identification or Lactococcus
 PT lactis and related species.
 XX
 PS Claim 6; SEQ ID NO 1794; 2504pp; French.
 XX
 CC The present invention is related to a Lactococcus lactis nucleotide
 CC sequence (ABA90521) and related proteins (ABB53300-ABB55621). The nucleic
 CC acid sequence is useful in the detection and/or amplification of nucleic
 CC acid sequence, particularly to identify Lactococcus lactis or related
 CC species. The proteins of the invention are useful for the biosynthesis or
 CC biodegradation of a composition of interest. The invention helps research
 CC in lactic bacteria, particularly useful in the production of yogurt and
 CC cheese. Note: The sequence data for this patent is based on equivalent
 CC patent WO200177334 (published 18-OCT-2001) which is available in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences. (Updated on 29-AUG-2003 to
 CC standardise OS field)
 XX
 SQ Sequence 162 AA;
 Query Match 79.3%; Score 23; DB 5; Length 162;
 Best Local Similarity 66.7%; Pred. No. 6.5e+02;

Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 MXVAF 7

Db 37 MTIAEF 42

Search completed: June 13, 2005, 13:56:04
Job time : 126 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2005, 13:35:12 ; Search time 30 seconds
(without alignments)
17.418 Million cell updates/sec

Title: 09730329-60ED
Perfect score: 29
Sequence: 1 VMXVAF 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 300 summaries

Database :

Issued Patents AA:*

1: /cgn2_6/ptodata/1/iaa/5A COMB.pep.*

2: /cgn2_6/ptodata/1/iaa/5B COMB.pep.*

3: /cgn2_6/ptodata/1/iaa/6A COMB.pep.*

4: /cgn2_6/ptodata/1/iaa/6B COMB.pep.*

5: /cgn2_6/ptodata/1/iaa/PCITUS COMB.pep.*

6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	27	93.1	7	4	US-09-724-566A-78
2	27	93.1	7	4	US-09-471-669A-78
3	27	93.1	8	4	US-09-724-566A-81
4	27	93.1	8	4	US-09-471-669A-81
5	27	93.1	68	4	US-09-393-634-80
6	27	93.1	450	4	US-09-543-681A-7260
7	25	86.2	194	4	US-09-134-000C-3800
8	24	82.8	188	4	US-09-902-540-14932
9	24	82.8	335	4	US-09-583-110-5099
10	24	82.8	620	4	US-08-637-670-40
11	24	82.8	927	3	US-09-134-001C-4831
12	24	82.8	972	3	US-08-335-844A-24
13	24	82.8	972	4	US-09-129-366-24
14	24	82.8	1176	4	US-09-489-039A-8879
15	23	79.3	54	2	US-08-456-647B-18
16	23	79.3	54	2	US-08-456-647B-46
17	23	79.3	54	2	US-08-456-647B-47
18	23	79.3	54	2	US-08-237-401A-18
19	23	79.3	54	2	US-08-237-401A-46
20	23	79.3	54	2	US-08-237-401A-47
21	23	79.3	74	4	US-09-248-796A-16636
22	23	79.3	90	4	US-09-393-634-76
23	23	79.3	132	4	US-09-134-000C-3642
24	23	79.3	193	4	US-09-270-767-41416
25	23	79.3	203	4	US-09-270-767-34950
26	23	79.3	203	4	US-09-270-767-50167
27	23	79.3	285	4	US-09-270-767-32289
28					Sequence 78, Appl
29					Sequence 78, Appl
30					Sequence 81, Appl
31					Sequence 81, Appl
32					Sequence 80, Appl
33					Sequence 7260, Ap
34					Sequence 3800, Ap
35					Sequence 14932, A
36					Sequence 5099, Ap
37					Sequence 40, Appl
38					Sequence 4831, Ap
39					Sequence 24, Appl
40					Sequence 24, Appl
41					Sequence 8879, Ap
42					Sequence 18, Appl
43					Sequence 46, Appl
44					Sequence 46, Appl
45					Sequence 18, Appl
46					Sequence 45, Appl
47					Sequence 47, Appl
48					Sequence 18, Appl
49					Sequence 46, Appl
50					Sequence 46, Appl
51					Sequence 18, Appl
52					Sequence 47, Appl
53					Sequence 16636, A
54					Sequence 76, Appl
55					Sequence 3642, Ap
56					Sequence 41416, A
57					Sequence 34950, A
58					Sequence 50167, A
59					Sequence 32289, A
60					
61					
62					
63					
64					
65					
66					
67					
68					
69					
70					
71					
72					
73					
74					
75					
76					
77					
78					
79					
80					
81					
82					
83					
84					
85					
86					
87					
88					
89					
90					
91					
92					
93					
94					
95					
96					
97					
98					
99					
100					

Sequence 47506, A

Sequence 13, Appl

Sequence 13, Appl

Sequence 31, Appl

Sequence 31, Appl

Sequence 9, Appl

Sequence 9, Appl

Sequence 6, Appl

Sequence 6, Appl

Sequence 8, Appl

Sequence 7, Appl

Sequence 7, Appl

Sequence 7, Appl

Sequence 7, Appl

Sequence 17, Appl

Sequence 16, Appl

Sequence 2, Appl

Sequence 2, Appl

Sequence 10, Appl

Sequence 10, Appl

Sequence 5567, Ap

Sequence 3, Appl

Sequence 3, Appl

Sequence 8, Appl

Sequence 8, Appl

Sequence 4, Appl

Sequence 4, Appl

Sequence 28422, A

Sequence 43247, A

Sequence 1818, Ap

Sequence 8, Appl

Sequence 22, Appl

Sequence 4257, Ap

Sequence 31250, A

Sequence 11, Appl

Sequence 7991, Ap

Sequence 10, Appl

Sequence 3, Appl

Sequence 6, Appl

Sequence 6, Appl

Sequence 5, Appl

Sequence 10, Appl

Sequence 4, Appl

Sequence 4, Appl

Sequence 6, Appl

Sequence 6, Appl

Sequence 7119, Ap

Sequence 6, Appl

Sequence 33, Appl

Sequence 3, Appl

Sequence 1, Appl

Sequence 10904, A

Sequence 2, Appl

Sequence 3, Appl

Sequence 3, Appl

Sequence 6, Appl

Sequence 1, Appl

Sequence 13, Appl

Sequence 13, Appl

Sequence 4, Appl

Sequence 4, Appl

Sequence 2, Appl

Sequence 4, Appl

Sequence 12, Appl

101	23	79.3	822	3	US-08-439-992A-2	Sequence 2, Appli	174	21	72.4	107	4	US-09-198-452A-982	Sequence 982, App
102	23	79.3	822	3	US-08-323-430-12	Sequence 12, Appl	175	21	72.4	112	4	US-09-583-110-4446	Sequence 4446, Ap
103	23	79.3	954	4	US-09-252-991A-32931	Sequence 32931, A	176	21	72.4	126	4	US-09-902-540-13031	Sequence 13031, A
104	23	79.3	973	1	US-08-162-809-8	Sequence 8, Appli	177	21	72.4	133	4	US-09-107-433-3161	Sequence 3161, Ap
105	23	79.3	2232	3	US-09-091-219-25	Sequence 25, Appl	178	21	72.4	135	4	US-09-248-796A-19769	Sequence 19769, A
106	23	79.3	2232	4	US-09-660-541-25	Sequence 25, Appl	179	21	72.4	149	3	US-08-836-075A-24	Sequence 24, Appl
107	23	79.3	2247	3	US-09-091-219-2	Sequence 2, Appli	180	21	72.4	149	4	US-09-270-767-48018	Sequence 48018, A
108	23	79.3	2247	3	US-09-660-541-2	Sequence 2, Appli	181	21	72.4	154	4	US-08-635-886C-236	Sequence 236, App
109	22	75.9	61	3	US-09-235-451-19	Sequence 19, Appl	182	21	72.4	154	4	US-08-974-690C-236	Sequence 236, App
110	22	75.9	61	3	US-09-978-303-19	Sequence 19, Appl	183	21	72.4	161	3	US-08-858-207A-284	Sequence 284, App
111	22	75.9	67	3	US-09-134-001C-4003	Sequence 4003, Ap	184	21	72.4	171	4	US-09-583-110-3868	Sequence 3868, Ap
112	22	75.9	116	4	US-09-489-039A-12512	Sequence 12512, A	185	21	72.4	172	4	US-09-107-532A-6805	Sequence 6805, Ap
113	22	75.9	173	4	US-09-248-796A-16707	Sequence 16707, A	186	21	72.4	179	4	US-09-634-238-296	Sequence 296, App
114	22	75.9	183	1	US-07-741-940-6	Sequence 6, Appli	187	21	72.4	181	4	US-09-107-433-4243	Sequence 4243, Ap
115	22	75.9	185	1	US-08-289-548A-6	Sequence 6, Appli	188	21	72.4	199	3	US-09-232-200-85	Sequence 85, Appl
116	22	75.9	185	1	US-08-452-654-6	Sequence 6, Appli	189	21	72.4	199	3	US-09-232-197-85	Sequence 85, Appl
117	22	75.9	185	1	US-08-452-655B-6	Sequence 6, Appli	190	21	72.4	199	3	US-09-232-201-85	Sequence 85, Appl
118	22	75.9	185	2	US-08-865-336-3	Sequence 3, Appli	191	21	72.4	199	4	US-09-232-195-85	Sequence 85, Appl
119	22	75.9	185	2	US-08-865-336-4	Sequence 4, Appli	192	21	72.4	200	4	US-09-543-681A-6662	Sequence 6662, Ap
120	22	75.9	185	3	US-08-450-582-6	Sequence 6, Appli	193	21	72.4	200	4	US-09-270-767-36343	Sequence 36343, A
121	22	75.9	185	4	US-08-449-731-6	Sequence 6, Appli	194	21	72.4	200	4	US-09-270-767-51560	Sequence 51560, A
122	22	75.9	210	1	US-07-741-940-4	Sequence 4, Appli	195	21	72.4	204	4	US-09-248-796A-15675	Sequence 15675, A
123	22	75.9	210	1	US-08-289-548A-4	Sequence 4, Appli	196	21	72.4	209	4	US-09-248-796A-17832	Sequence 17832, A
124	22	75.9	210	1	US-08-452-654-4	Sequence 4, Appli	197	21	72.4	212	4	US-09-543-681A-6021	Sequence 6021, Ap
125	22	75.9	210	1	US-08-452-655B-4	Sequence 4, Appli	198	21	72.4	234	4	US-09-107-532A-6523	Sequence 6523, Ap
126	22	75.9	210	3	US-08-450-582-4	Sequence 4, Appli	199	21	72.4	238	3	US-09-605-858-35	Sequence 35, Appl
127	22	75.9	210	4	US-08-449-731-4	Sequence 4, Appli	200	21	72.4	239	4	US-09-489-039A-11136	Sequence 11136, A
128	22	75.9	221	4	US-09-949-016-10154	Sequence 10154, A	201	21	72.4	243	4	US-09-248-796A-23192	Sequence 23192, A
129	22	75.9	236	3	US-09-605-858-34	Sequence 34, Appl	202	21	72.4	249	3	US-09-154-083-2	Sequence 2, Appli
130	22	75.9	252	4	US-09-248-796A-26937	Sequence 26937, A	203	21	72.4	250	4	US-09-107-433-2880	Sequence 2880, Ap
131	22	75.9	275	4	US-09-079-030-224	Sequence 224, App	204	21	72.4	253	4	US-09-107-532A-5072	Sequence 5072, Ap
132	22	75.9	320	4	US-09-543-681A-5983	Sequence 5983, Ap	205	21	72.4	253	4	US-09-583-110-3446	Sequence 3446, Ap
133	22	75.9	379	4	US-09-902-540-13539	Sequence 13539, A	206	21	72.4	254	4	US-09-198-452A-632	Sequence 632, App
134	22	75.9	392	4	US-09-107-532A-4840	Sequence 4840, Ap	254	21	72.4	254	4	US-09-438-185A-591	Sequence 591, App
135	22	75.9	397	4	US-09-583-110-4402	Sequence 4402, Ap	208	21	72.4	255	4	US-09-134-000C-6258	Sequence 6258, Ap
136	22	75.9	400	4	US-09-107-433-4600	Sequence 4600, Ap	210	21	72.4	258	4	US-09-107-433-47938	Sequence 4798, Ap
137	22	75.9	437	4	US-09-710-279-2960	Sequence 2960, Ap	210	21	72.4	263	4	US-09-902-540-9734	Sequence 9734, Ap
138	22	75.9	442	4	US-09-489-039A-13610	Sequence 13610, A	211	21	72.4	271	4	US-09-107-532A-5071	Sequence 5071, Ap
139	22	75.9	445	4	US-09-252-991A-28986	Sequence 28986, A	212	21	72.4	274	4	US-09-270-767-45001	Sequence 45001, A
140	22	75.9	463	3	US-09-134-001C-4160	Sequence 4160, Ap	213	21	72.4	274	4	US-09-248-796A-20041	Sequence 20041, A
141	22	75.9	489	4	US-09-252-991A-22896	Sequence 22896, A	214	21	72.4	279	2	US-08-701-191A-23	Sequence 23, Appl
142	22	75.9	497	4	US-09-902-540-10388	Sequence 10388, A	215	21	72.4	279	4	US-09-664-526-23	Sequence 23, Appl
143	22	75.9	511	4	US-09-134-000C-5362	Sequence 5362, Ap	215	21	72.4	279	4	US-09-664-526-23	Sequence 32518, A
144	22	75.9	605	2	US-08-687-956A-1	Sequence 1, Appli	217	21	72.4	279	4	US-09-270-767-47735	Sequence 47735, A
145	22	75.9	605	3	US-09-134-001C-4425	Sequence 4425, Ap	218	21	72.4	283	4	US-09-583-110-3146	Sequence 3146, Ap
146	22	75.9	784	4	US-09-252-991A-32809	Sequence 32809, A	219	21	72.4	288	4	US-09-252-991A-18721	Sequence 18721, A
147	22	75.9	945	4	US-09-949-016-8172	Sequence 8172, Ap	220	21	72.4	294	4	US-09-523-263B-20	Sequence 20, Appl
148	22	75.9	951	1	US-08-162-809-2	Sequence 2, Appli	221	21	72.4	294	4	US-09-248-796A-16415	Sequence 16415, A
149	22	75.9	970	2	US-08-449-645A-11	Sequence 11, Appl	222	21	72.4	295	4	US-09-543-681A-5660	Sequence 5660, Ap
150	22	75.9	970	2	US-08-702-367A-11	Sequence 11, Appl	223	21	72.4	311	4	US-09-583-110-3705	Sequence 3705, Ap
151	22	75.9	970	5	PCT-US95-04681-11	Sequence 11, Appl	224	21	72.4	312	4	US-09-543-681A-7100	Sequence 7100, Ap
152	22	75.9	973	1	US-08-162-809-10	Sequence 10, Appl	225	21	72.4	315	4	US-09-107-433-4817	Sequence 4817, Ap
153	22	75.9	984	2	US-08-673-789-6	Sequence 6, Appli	226	21	72.4	323	4	US-09-540-236-2794	Sequence 2794, Ap
154	22	75.9	984	4	US-09-949-016-6502	Sequence 6502, Ap	227	21	72.4	325	4	US-09-248-796A-14430	Sequence 14430, A
155	22	75.9	988	1	US-08-162-809-14	Sequence 14, Appl	228	21	72.4	336	1	US-07-904-073-2	Sequence 2, Appli
156	22	75.9	993	3	US-08-368-776A-11	Sequence 11, Appl	229	21	72.4	336	1	US-07-904-071-2	Sequence 2, Appli
157	22	75.9	994	3	US-08-542-635-2	Sequence 2, Appli	230	21	72.4	336	1	US-08-442-043A-16	Sequence 16, Appl
158	22	75.9	994	3	US-08-368-776A-12	Sequence 12, Appl	231	21	72.4	336	4	US-08-441-893A-16	Sequence 16, Appl
159	22	75.9	995	1	US-08-162-809-18	Sequence 18, Appl	232	21	72.4	336	4	US-09-252-991A-24121	Sequence 24121, A
160	22	75.9	998	2	US-08-449-645A-17	Sequence 17, Appl	233	21	72.4	342	4	US-09-489-039A-8158	Sequence 8158, Ap
161	22	75.9	998	2	US-08-702-367A-17	Sequence 17, Appl	234	21	72.4	346	4	US-09-489-039A-7365	Sequence 7365, Ap
162	22	75.9	998	3	US-08-368-776A-2	Sequence 2, Appli	235	21	72.4	350	4	US-09-489-039A-9140	Sequence 9140, Ap
163	22	75.9	998	5	PCT-US95-04681-17	Sequence 17, Appl	236	21	72.4	353	4	US-09-252-991A-17097	Sequence 17097, A
164	22	75.9	998	5	PCT-US95-04681-17	Sequence 17, Appl	237	21	72.4	356	4	US-09-328-352-7671	Sequence 7671, A
165	22	75.9	998	5	PCT-US96-00419-2	Sequence 2, Appli	238	21	72.4	360	4	US-09-248-796A-18247	Sequence 18247, A
166	22	75.9	1005	4	US-09-949-016-9901	Sequence 9901, Ap	239	21	72.4	370	4	US-09-543-681A-6187	Sequence 6187, Ap
167	22	75.9	1007	3	US-08-961-083-216	Sequence 216, App	240	21	72.4	371	4	US-09-354-123-10	Sequence 10, Appl
168	22	75.9	1007	4	US-09-536-784-216	Sequence 216, App	241	21	72.4	380	4	US-09-133-000C-5772	Sequence 5772, Ap
169	22	75.9	1011	1	US-08-162-809-12	Sequence 12, Appl	242	21	72.4	390	4	US-09-933-313B-8	Sequence 8, Appli
170	22	75.9	1033	4	US-09-328-352-5138	Sequence 5138, Ap	243	21	72.4	397	4	US-09-328-352-7457	Sequence 7457, Ap
171	22	75.9	1529	4	US-09-107-433-4771	Sequence 4771, Ap	244	21	72.4	404	3	US-09-242-859A-11	Sequence 11, Appl
172	22	75.9	2138	4	US-09-583-110-5274	Sequence 5274, Ap	245	21	72.4	404	4	US-09-328-352-6361	Sequence 6361, Ap
173	21	72.4	72	4	US-09-248-796A-23301	Sequence 23301, A	246	21	72.4	405	4	US-09-538-092-1059	Sequence 1059, Ap

247 21 72.4 416 3 US-09-355-115-9 Sequence 9, Appli
248 21 72.4 416 4 US-09-915-182-2 Sequence 2, Appli
249 21 72.4 416 4 US-09-591-279A-45 Sequence 45, Appli
250 21 72.4 416 4 US-09-949-016-7485 Sequence 7485, Ap
251 21 72.4 421 4 US-09-540-236-2896 Sequence 2896, Ap
252 21 72.4 429 4 US-09-328-352-4875 Sequence 4875, Ap
253 21 72.4 437 4 US-09-252-991A-25332 Sequence 25332, A
254 21 72.4 432 4 US-09-902-540-10477 Sequence 10477, A
255 21 72.4 469 4 US-09-543-681A-5423 Sequence 5423, A
256 21 72.4 475 4 US-09-328-352-5697 Sequence 5697, Ap
257 21 72.4 482 4 US-09-489-039A-9369 Sequence 9369, Ap
258 21 72.4 510 4 US-09-489-039A-11123 Sequence 11123, A
259 21 72.4 517 4 US-09-538-092-851 Sequence 851, App
260 21 72.4 517 4 US-09-538-092-10477 Sequence 1047, Ap
261 21 72.4 520 4 US-09-949-016-10653 Sequence 10653, A
262 21 72.4 522 4 US-09-252-991A-30451 Sequence 30451, A
263 21 72.4 534 4 US-09-489-039A-8550 Sequence 8550, Ap
264 21 72.4 543 4 US-09-248-796A-19658 Sequence 19658, A
265 21 72.4 559 3 US-09-242-690A-15 Sequence 15, Appli
266 21 72.4 559 3 US-09-298-924-6 Sequence 6, Appli
267 21 72.4 559 4 US-09-908-855-15 Sequence 15, Appli
268 21 72.4 569 1 US-07-821-716-2 Sequence 2, Appli
269 21 72.4 569 2 US-08-381-603-2 Sequence 2, Appli
270 21 72.4 569 3 US-08-924-376-2 Sequence 2, Appli
271 21 72.4 569 3 US-08-685-212-2 Sequence 2, Appli
272 21 72.4 569 3 US-09-173-151A-31 Sequence 31, Appli
273 21 72.4 569 4 US-08-466-932A-2 Sequence 2, Appli
274 21 72.4 569 4 US-08-406-824A-6 Sequence 6, Appli
275 21 72.4 569 4 US-09-949-016-6000 Sequence 6000, Ap
276 21 72.4 569 5 PCT-US94-02414-2 Sequence 2, Appli
277 21 72.4 569 5 PCT-US96-08899-2 Sequence 2, Appli
278 21 72.4 581 4 US-09-538-092-383 Sequence 383, App
279 21 72.4 588 4 US-09-949-016-8572 Sequence 8572, Ap
280 21 72.4 607 4 US-09-248-796A-17281 Sequence 17281, A
281 21 72.4 645 4 US-09-252-991A-22095 Sequence 22095, A
282 21 72.4 649 4 US-09-538-092-1357 Sequence 1357, Ap
283 21 72.4 653 4 US-09-252-991A-27932 Sequence 27932, A
284 21 72.4 711 3 US-08-772-270A-12 Sequence 12, Appli
285 21 72.4 713 4 US-09-949-016-10902 Sequence 10902, A
286 21 72.4 714 4 US-09-902-540-15419 Sequence 15419, A
287 21 72.4 718 4 US-09-252-991A-33109 Sequence 33109, A
288 21 72.4 729 4 US-09-949-016-6988 Sequence 6988, Ap
289 21 72.4 733 4 US-09-949-016-11165 Sequence 11165, A
290 21 72.4 773 4 US-09-913-301-5 Sequence 5, Appli
291 21 72.4 801 4 US-09-913-301-7 Sequence 7, Appli
292 21 72.4 804 4 US-09-913-301-2 Sequence 2, Appli
293 21 72.4 863 4 US-09-902-540-10725 Sequence 10725, A
294 21 72.4 908 4 US-09-252-991A-24682 Sequence 24682, A
295 21 72.4 910 4 US-09-313-942-28 Sequence 28, Appli
296 21 72.4 926 3 US-08-855-910-2 Sequence 2, Appli
297 21 72.4 984 2 US-08-673-789-9 Sequence 9, Appli
298 21 72.4 984 2 US-08-449-645A-19 Sequence 19, Appli
299 21 72.4 984 2 US-08-702-367A-19 Sequence 19, Appli
300 21 72.4 984 5 PCT-US95-04681-19 Sequence 19, Appli

ALIGNMENTS

RESULT 1
US-09-724-566A-78
; Sequence 78, Application US/09724566A
; Patent No. 6627739
; GENERAL INFORMATION:
; APPLICANT: Anderson, John P.
; APPLICANT: Basi, Gurigbal
; APPLICANT: Doane, Minh Tam
; APPLICANT: Frigon, No. 6627739mand
; APPLICANT: John, Varghese
; APPLICANT: Power, Michael
; APPLICANT: Sinha, Sukanto
; APPLICANT: Tatsuno, Gwen
; APPLICANT: Tung, Jay

; APPLICANT: Wang, Shuwen
; APPLICANT: McConlogue, Lisa
; TITLE OF INVENTION: Beta-Secretase Enzyme Compositions and
; TITLE OF INVENTION: Methods
; FILE REFERENCE: 228-US-NEWC2
; CURRENT APPLICATION NUMBER: US/09/724,566A
; CURRENT FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: US 09/501,708
; PRIOR FILING DATE: 2000-02-10
; PRIOR APPLICATION NUMBER: 60/119,571
; PRIOR FILING DATE: 1999-02-10
; PRIOR APPLICATION NUMBER: 60/139,172
; PRIOR FILING DATE: 1999-06-15
; NUMBER OF SEQ ID NOS: 104
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 78
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Peptide inhibitor P3-P4' XD-V
; NAME/KEY: MOD_RES
; LOCATION: 3
; OTHER INFORMATION: Xaa is hydroxyethylene or statine
US-09-724-566A-78

Query Match 93.1%; Score 27; DB 4: Length 7;
Best Local Similarity 100.0%; Pred. No. 4.1e+05; Indels 0;
Matches 7; Conservative 0; Mismatches 0; Gaps 0;

OY 1 VMXVAEF 7
| | | | |
Db 1 VMXVAEF 7

RESULT 2
US-09-471-669A-78
; Sequence 78, Application US/09471669A
; Patent No. 6830918
; GENERAL INFORMATION:
; APPLICANT: Anderson, John P.
; APPLICANT: Basi, Gurigbal
; APPLICANT: Doane, Minh Tam
; APPLICANT: Frigon, No. 6830918mand
; APPLICANT: John, Varghese
; APPLICANT: Power, Michael
; APPLICANT: Sinha, Sukanto
; APPLICANT: Tatsuno, Gwen
; APPLICANT: Tung, Jay
; APPLICANT: Wang, Shuwen
; APPLICANT: McConlogue, Lisa
; APPLICANT: Elan Pharmaceuticals, Inc.
; TITLE OF INVENTION: BETA-SECRETASE ENZYME COMPOSITIONS AND METHODS
; FILE REFERENCE: 015270-006430US
; CURRENT APPLICATION NUMBER: US/09/471,669A
; CURRENT FILING DATE: 1999-12-24
; PRIOR APPLICATION NUMBER: US 60/114,408
; PRIOR FILING DATE: 1998-12-31
; PRIOR APPLICATION NUMBER: US 60/119,571
; PRIOR FILING DATE: 1999-02-10
; PRIOR APPLICATION NUMBER: US 60/139,172
; PRIOR FILING DATE: 1999-06-15
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 78
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Peptide
; OTHER INFORMATION: Inhibitor P3-P4' XD-V
; FEATURE:

```
; NAME/KEY: MOD_RES
; LOCATION: (3)
; OTHER INFORMATION: Xaa is hydroxyethylene or statine
US-09-471-669A-78
```

```
Query Match          93.1%; Score 27; DB 4; Length 7;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches              7; Conservative 0; Mismatches 0; Indels
```

```
0; Gaps 0;
```

```
Qy 1 VMXVAEF 7
    |||||
Db 1 VMXVAEF 7
```

RESULT 3

```
US-09-724-566A-81
; Sequence 81, Application US/09724566A
; Patent No. 6627739
; GENERAL INFORMATION:
; APPLICANT: Anderson, John P.
; APPLICANT: Basi, Guribbal
; APPLICANT: Doane, Minh Tam
; APPLICANT: Frigon, No. 6627739mand
; APPLICANT: John, Varghese
; APPLICANT: Power, Michael
; APPLICANT: Sinha, Sukanto
; APPLICANT: Tatsuno, Gwen
; APPLICANT: Tung, Jay
; APPLICANT: Wang, Shuwen
; APPLICANT: McConlogue, Lisa
; TITLE OF INVENTION: Beta-Secretase Enzyme Compositions and
; FILE REFERENCE: 228-US-NEWC2
; CURRENT APPLICATION NUMBER: US/09/724,566A
; CURRENT FILING DATE: 2000-11-28
; PRIOR FILING DATE: 2000-02-10
; PRIOR APPLICATION NUMBER: 60/119,571
; PRIOR FILING DATE: 1999-02-10
; PRIOR APPLICATION NUMBER: 60/139,172
; PRIOR FILING DATE: 1999-06-15
; NUMBER OF SEQ ID NOS: 104
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 81
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Peptide inhibitor P4-P4' XD-V
; NAME/KEY: MOD_RES
; LOCATION: 4
; OTHER INFORMATION: Xaa is hydroxyethylene or statine
US-09-724-566A-81
```

```
Query Match          93.1%; Score 27; DB 4; Length 8;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches              7; Conservative 0; Mismatches 0; Indels
```

```
0; Gaps 0;
```

```
Qy 1 VMXVAEF 7
    |||||
Db 2 VMXVAEF 8
```

RESULT 4

```
US-09-471-669A-81
; Sequence 81, Application US/09471669A
; Patent No. 6830918
; GENERAL INFORMATION:
; APPLICANT: Anderson, John P.
; APPLICANT: Basi, Guribbal
; APPLICANT: Doane, Minh Tam
; APPLICANT: Frigon, No. 6830918mand
```

```
; APPLICANT: John, Varghese
; APPLICANT: Power, Michael
; APPLICANT: Sinha, Sukanto
; APPLICANT: Tatsuno, Gwen
; APPLICANT: Tung, Jay
; APPLICANT: Wang, Shuwen
; APPLICANT: McConlogue, Lisa
; TITLE OF INVENTION: BETA-SECRETASE ENZYME COMPOSITIONS AND METHODS
; FILE REFERENCE: 015270-0064300S
; CURRENT APPLICATION NUMBER: US/09/471,669A
; CURRENT FILING DATE: 1999-12-24
; PRIOR APPLICATION NUMBER: US 60/114,408
; PRIOR FILING DATE: 1998-12-31
; PRIOR APPLICATION NUMBER: US 60/119,571
; PRIOR FILING DATE: 1999-02-10
; PRIOR APPLICATION NUMBER: US 60/139,172
; PRIOR FILING DATE: 1999-06-15
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 81
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Peptide
; OTHER INFORMATION: inhibitor P4-P4' XD-V
; NAME/KEY: MOD_RES
; LOCATION: (4)
; OTHER INFORMATION: Xaa is hydroxyethylene or statine.
US-09-471-669A-81
```

```
Query Match          93.1%; Score 27; DB 4; Length 8;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches              7; Conservative 0; Mismatches 0; Indels
```

```
0; Gaps 0;
```

```
Qy 1 VMXVAEF 7
    |||||
Db 2 VMXVAEF 8
```

RESULT 5

```
US-09-393-634-80
; Sequence 80, Application US/09393634
; Patent No. 6558910
; GENERAL INFORMATION:
; APPLICANT: Zuker, Charles S.
; APPLICANT: Adler, Jon Elliot
; APPLICANT: Ryba, Nick
; APPLICANT: Mueller, Ken
; APPLICANT: Hoon, Mark
; APPLICANT: The Regents of the University of California
; APPLICANT: The Government of the United States of America
; APPLICANT: as represented by the Secretary of the
; APPLICANT: Department of Health and Human Services
; TITLE OF INVENTION: SF, a No. 6558910el Family of Taste Receptors
; FILE REFERENCE: 02307E-098000US
; CURRENT APPLICATION NUMBER: US/09/393,634
; CURRENT FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 92
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 80
; LENGTH: 68
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: human GR24
US-09-393-634-80
```

```
Query Match          93.1%; Score 27; DB 4; Length 68;
Best Local Similarity 71.4%; Pred. No. 9.2;
Matches              5; Conservative 1; Mismatches 1; Indels
```

```
0; Gaps 0;
```



```
Qy 1 VMXVAEF 7
Db 11 IMVAEF 17

RESULT 6
US-09-543-681A-7260
; Sequence 7260, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 7260
; LENGTH: 450
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-7260

Query Match 93.1%; Score 27; DB 4; Length 450;
Best Local Similarity 85.7%; Pred. No. 69;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VMXVAEF 7
Db 286 VMVVAEF 292

RESULT 7
US-09-134-000C-3800
; Sequence 3800, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3800
; LENGTH: 194
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-134-000C-3800

Query Match 86.2%; Score 25; DB 4; Length 194;
Best Local Similarity 71.4%; Pred. No. 89;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VMXVAEF 7
Db 112 VLSVAEF 118

RESULT 8
US-09-902-540-14932
; Sequence 14932, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
```

```
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 14932
; LENGTH: 188
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-14932

Query Match 82.8%; Score 24; DB 4; Length 188;
Best Local Similarity 83.3%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 MXVAEF 7
Db 1 MTVAEF 6

RESULT 9
US-09-583-110-5099
; Sequence 5099, Application US/09583110
; Patent No. 6699703
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
; FILE REFERENCE: PATH00-07A
; CURRENT APPLICATION NUMBER: US/09/583,110
; CURRENT FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/107,433
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/085,131
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: US 60/051,553
; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 5322
; SEQ ID NO 5099
; LENGTH: 335
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-583-110-5099

Query Match 82.8%; Score 24; DB 4; Length 335;
Best Local Similarity 71.4%; Pred. No. 2.8e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 VMXVAEF 7
Db 224 VMSAEF 230

RESULT 10
US-08-637-670-40
; Sequence 40, Application US/08637670
; Patent No. 6413521
; GENERAL INFORMATION:
; APPLICANT: MCMICHAEL-PHILLIPS et al.
; TITLE OF INVENTION: Helminth Parasite Antigen with
; TITLE OF INVENTION: Aminopeptidase-like Activity
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Barbara G. Ernst
; STREET: 555 13TH STREET, NW Suite 701E
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
```

```
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/637,670
; FILING DATE: 26-JUN-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Ernst, Barbara G.
; REGISTRATION NUMBER: 30,377
; REFERENCE/DOCKET NUMBER: 1811-0232
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-783-6040
; TELEFAX: 202-783-6031
; INFORMATION FOR SEQ ID NO: 40:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 620 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-637-670-40

Query Match 82.8%; Score 24; DB 4; Length 620;
Best Local Similarity 71.4%; Pred. No. 5.5e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
Db 203 VISVAEP 209

RESULT 11
US-09-134-001C-4831
; Sequence 4831, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 4831
; LENGTH: 927
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
; US-09-134-001C-4831

Query Match 82.8%; Score 24; DB 3; Length 927;
Best Local Similarity 83.3%; Pred. No. 8.4e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 MXVAEF 7
Db 126 MSVAEP 131

RESULT 12
US-08-335-844A-24
; Sequence 24, Application US/08335844A
; Patent No. 6068503
; GENERAL INFORMATION:
; APPLICANT: GRAHAM, MARGARET
; APPLICANT: SMITH, TREVOR STANLEY
; APPLICANT: MUNN, EDWARD ALBERT
; APPLICANT: KNOX, DAVID PATRICK

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/637,670
; FILING DATE: 26-JUN-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Ernst, Barbara G.
; REGISTRATION NUMBER: 30,377
; REFERENCE/DOCKET NUMBER: 1811-0232
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-783-6040
; TELEFAX: 202-783-6031
; INFORMATION FOR SEQ ID NO: 40:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 620 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-637-670-40

; APPLICANT: OLIVER, JOANNA JANE
; APPLICANT: NEWTON, SUSAN ELIZABETH
; TITLE OF INVENTION: RECOMBINANT DNA MOLECULES ENCODING
; TITLE OF INVENTION: AMINOPEPTIDASE ENZYMES AND THEIR USE IN THE PREPARATION OF
; TITLE OF INVENTION: VACCINES AGAINST HELMINTH INFECTIONS
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rothwell, Figg, Ernst & Kurz
; STREET: Suite 701-B, 555 Thirteenth St., N.W
; CITY: Washington
; STATE: D. C.
; COUNTRY: U.S.A.
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/335,844A
; FILING DATE: 09-JAN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB PCT/GB93/00943
; FILING DATE: 06-MAY-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9209936
; FILING DATE: 08-MAY-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: WALKER, Barbara W.
; REGISTRATION NUMBER: 35,400
; REFERENCE/DOCKET NUMBER: 1181-223A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)783-6040
; TELEFAX: (202)783-6031
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 972 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-335-844A-24

Query Match 82.8%; Score 24; DB 3; Length 972;
Best Local Similarity 71.4%; Pred. No. 8.8e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
Db 542 VISVAEP 548

RESULT 13
US-09-129-366-24
; Sequence 24, Application US/09129366
; Patent No. 6534638
; GENERAL INFORMATION:
; APPLICANT: GRAHAM, MARGARET
; APPLICANT: SMITH, TREVOR STANLEY
; APPLICANT: MUNN, EDWARD ALBERT
; APPLICANT: KNOX, DAVID PATRICK
; APPLICANT: OLIVER, JOANNA JANE
; APPLICANT: NEWTON, SUSAN ELIZABETH
; TITLE OF INVENTION: RECOMBINANT DNA MOLECULES ENCODING
; TITLE OF INVENTION: AMINOPEPTIDASE ENZYMES AND THEIR USE IN THE PREPARATION OF
; TITLE OF INVENTION: VACCINES AGAINST HELMINTH INFECTIONS
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rothwell, Figg, Ernst & Kurz
; STREET: Suite 701-B, 555 Thirteenth St., N.W
; CITY: Washington
; STATE: D. C.
; COUNTRY: U.S.A.
```

ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
FILING DATE: 05-AUG-1998
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/335,844
FILING DATE: 09-JAN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB PCT/GB93/00943
FILING DATE: 06-MAY-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9209936
FILING DATE: 08-MAY-1992
ATTORNEY/AGENT INFORMATION:
NAME: ERNST, Barbara G.
REGISTRATION NUMBER: 30,377
REFERENCE/DOCKET NUMBER: 1181-241A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)783-6040
TELEFAX: (202)783-6031
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 972 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-129-366-24

Query Match 82.8%; Score 24; DB 4; Length 972;
Best Local Similarity 71.4%; Pred. No. 8.e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
|: |||||
DB 542 VISVAEF 548

RESULT 14

US-09-489-039A-8879
Sequence 8879, Application US/09489039A
Patent No. 6610836
GENERAL INFORMATION:
APPLICANT: Gary Breton et. al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
FILE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 2709.2004001
CURRENT APPLICATION NUMBER: US/09/489,039A
CURRENT FILING DATE: 2000-01-27
PRIOR APPLICATION NUMBER: US 60/117,747
PRIOR FILING DATE: 1999-01-29
NUMBER OF SEQ ID NOS: 14342
SEQ ID NO 8879
LENGTH: 1176
TYPE: PRT
ORGANISM: Klebsiella pneumoniae
US-09-489-039A-8879

Query Match 82.8%; Score 24; DB 4; Length 1176;
Best Local Similarity 71.4%; Pred. No. 1.1.e+03;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
|: |||||
DB 779 VMKVAQF 785

RESULT 15

US-08-456-647B-18
Sequence 18, Application US/08456647B
Patent No. 5811516
GENERAL INFORMATION:
APPLICANT: Lemke Ph.D. et al., Greg E.
TITLE OF INVENTION: PROTEIN-TYROSINE KINASE GENES
NUMBER OF SEQUENCES: 54
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: CA
COUNTRY: US
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/456,647B
FILING DATE: 02-JUN-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/237,401
FILING DATE: 02-MAY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/884,486
FILING DATE: 15-MAY-1992
ATTORNEY/AGENT INFORMATION:
NAME: Wetherell Ph.D., John R.
REGISTRATION NUMBER: 31,678
REFERENCE/DOCKET NUMBER: 07251/007002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 678-5070
TELEFAX: (619) 678-5099
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 54 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-456-647B-18

Query Match 79.3%; Score 23; DB 2; Length 54;
Best Local Similarity 57.1%; Pred. No. 72;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
|: ||:|
DB 9 VMKIADF 15

RESULT 16

US-08-456-647B-46
Sequence 46, Application US/08456647B
Patent No. 5811516
GENERAL INFORMATION:
APPLICANT: Lemke Ph.D. et al., Greg E.
TITLE OF INVENTION: PROTEIN-TYROSINE KINASE GENES
NUMBER OF SEQUENCES: 54
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: CA
COUNTRY: US
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25

```
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/456,647B
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/237,401
; FILING DATE: 02-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/884,486
; FILING DATE: 15-MAY-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Wetherell Ph.D., John R.
; REGISTRATION NUMBER: 31,678
; REFERENCE/DOCKET NUMBER: 07251/007002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 678-5070
; TELEFAX: (619) 678-5099
; INFORMATION FOR SEQ ID NO: 46:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 54 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-456-647B-46

Query Match          79.3%; Score 23; DB 2; Length 54;
Best Local Similarity 57.1%; Pred. No. 72;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy      1 VMXVAEF 7
Db      9 VMKIADF 15

RESULT 17
US-08-456-647B-47
; Sequence 47, Application US/08456647B
; Patent No. 5811516
; GENERAL INFORMATION:
; APPLICANT: Lemke Ph.D. et al., Greg E.
; TITLE OF INVENTION: PROTEIN-TYROSINE KINASE GENES
; NUMBER OF SEQUENCES: 54
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: US
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/456,647B
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/237,401
; FILING DATE: 02-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/884,486
; FILING DATE: 15-MAY-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Wetherell Ph.D., John R.
; REGISTRATION NUMBER: 31,678
; REFERENCE/DOCKET NUMBER: 07251/007002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 678-5070
; TELEFAX: (619) 678-5099
; INFORMATION FOR SEQ ID NO: 47:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 54 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-456-647B-47
```

```
; SEQUENCE CHARACTERISTICS:
; LENGTH: 54 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-456-647B-47

Query Match          79.3%; Score 23; DB 2; Length 54;
Best Local Similarity 57.1%; Pred. No. 72;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy      1 VMXVAEF 7
Db      9 VMKIADF 15

RESULT 18
US-08-237-401A-18
; Sequence 18, Application US/08237401A
; Patent No. 5837448
; GENERAL INFORMATION:
; APPLICANT: Lemke Ph.D. et al., Greg E.
; TITLE OF INVENTION: PROTEIN-TYROSINE KINASE GENES
; NUMBER OF SEQUENCES: 54
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: US
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/237,401A
; FILING DATE: 02-MAY-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/884,486
; FILING DATE: 15-MAY-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile Ph.D., Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 07251/007001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 678-5070
; TELEFAX: (619) 678-5099
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 54 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-237-401A-18

Query Match          79.3%; Score 23; DB 2; Length 54;
Best Local Similarity 57.1%; Pred. No. 72;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy      1 VMXVAEF 7
Db      9 VMKIADF 15

RESULT 19
US-08-237-401A-46
; Sequence 46, Application US/08237401A
; Patent No. 5837448
; GENERAL INFORMATION:
```

; APPLICANT: Lemke Ph.D. et al., Greg E.
; TITLE OF INVENTION: PROTEIN-TYROSINE KINASE GENES
; NUMBER OF SEQUENCES: 54
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: US
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/237,401A
; FILING DATE: 02-MAY-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/884,486
; FILING DATE: 15-MAY-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile Ph.D., Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 07251/007001
; TELEPHONE: (619) 678-5070
; TELEFAX: (619) 678-5099
; INFORMATION FOR SEQ ID NO: 46:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 54 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-237-401A-46

Query Match 79.3%; Score 23; DB 2; Length 54;
Best Local Similarity 57.1%; Pred. No. 72;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
Db 9 VMKIADF 15

RESULT 20
US-08-237-401A-47
; Sequence 47, Application US/08237401A
; Patent No. 5837448
; GENERAL INFORMATION:
; APPLICANT: Lemke Ph.D. et al., Greg E.
; TITLE OF INVENTION: PROTEIN-TYROSINE KINASE GENES
; NUMBER OF SEQUENCES: 54
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: US
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/237,401A
; FILING DATE: 02-MAY-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/884,486

; FILING DATE: 15-MAY-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile Ph.D., Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 07251/007001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 678-5070
; TELEFAX: (619) 678-5099
; INFORMATION FOR SEQ ID NO: 47:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 54 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-237-401A-47

Query Match 79.3%; Score 23; DB 2; Length 54;
Best Local Similarity 57.1%; Pred. No. 72;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
Db 9 VMKIADF 15

RESULT 21
US-09-248-796A-16636
; Sequence 16636, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 16636
; LENGTH: 74
; TYPE: PRT
; ORGANISM: Candida albicans
; US-09-248-796A-16636

Query Match 79.3%; Score 23; DB 4; Length 74;
Best Local Similarity 71.4%; Pred. No. 1e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
Db 35 VMKAAEF 41

RESULT 22
US-09-393-634-76
; Sequence 76, Application US/09393634
; Patent No. 6558910
; GENERAL INFORMATION:
; APPLICANT: Zuker, Charles S.
; APPLICANT: Adler, Jon Elliot
; APPLICANT: Ryba, Nick
; APPLICANT: Mueller, Ken
; APPLICANT: Hoon, Mark
; APPLICANT: The Regents of the University of California
; APPLICANT: The Government of the United States of America
; APPLICANT: as represented by the Secretary of the
; APPLICANT: Department of Health and Human Services
; TITLE OF INVENTION: SF, a No. 6558910el Family of Taste Receptors
; FILE REFERENCE: 02307E-098000US

```

; CURRENT APPLICATION NUMBER: US/09/393,634
; CURRENT FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 92
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 76
; LENGTH: 90
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: human GR22
; NAME/KEY: MOD_RES
; LOCATION: (1)..(90)
; OTHER INFORMATION: xaa = any amino acid
; US-09-393-634-76

```

Query Match 79.3%; Score 23; DB 4; Length 90;
Best Local Similarity 42.9%; Pred. No. 1.2e+02;
Matches 3; Conservative 3; Mismatches 1; Indels

Qy 1 VMXVAEF 7
::: |||
Dp 11 ILAIAEF 1

```

RESULT 23
US-09-134-000C-3642
; Sequence 3642, Application US/09134000C
; Patent NO. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND
; TITLE OF INVENTION: ENTEROCOCCUS FAEC
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09134,
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,77
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3642
; LENGTH: 132
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-134-000C-3642

```

```
Query Match      79.3%; Score 23; DB 4; Length 132;
Best Local Similarity 57.1%; Pred.No. 1.9e+02;
Matches 4; Conservative 2; Mismatches 1; Indels
```

Qy	1	VMXVAEF	7
		:	:
		:	:
D _b	9	LM DIAEF	15

```

RESULT 24
US-09-270-767-41416
/ SEQUENCE 41416, Application US/09270767
/ Patent No. 6703491
/ GENERAL INFORMATION:
/ APPLICANT: Homburger et al.
/ TITLE OF INVENTION: Nucleic acids and
/ FILE REFERENCE: File Reference: 7326-0
/ CURRENT APPLICATION NUMBER: US/09270,
/ CURRENT FILING DATE: 1999-03-17
/ NUMBER OF SEQ ID NOS: 62517
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 41416
/ LENGTH: 193
/ TYPE: PRT
/ ORGANISM: Drosophila melanogaster
US-09-270-767-41416

```

Query Match 79.3%; Score 23; DB 4; Length 193;

Best Local Similarity	57.1%;	Pred. No. 2.8e+02;
Matches	4; Conservative	2; Mismatches 1; Indels
		Gaps 0; Gaps 0;

QY 1 VMXVAEF 7
:| | :|
Dp 99 IMPVAEF 105

```

RESULT 25
US-09-270-767-34950
; Sequence 34950, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 34950
; LENGTH: 203
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-34950

```

Query Match 79.3%; Score 23; DB 4; Length 203;
Best Local Similarity 83.3%; Pred. No. 2.9e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy	2	MXVAEF	7
D _b	16	MHVAEF	21

```

RESULT 26
US-09-270-767-50167
; Sequence 50167, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 50167
; LENGTH: 203
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-50167

```

Query Match	79.3%;	Score 23;	DB 4;	Length 203;
Best Local Similarity	83.3%;	Pred. No. 2.9e+02;		
Matches 5;	Conservative	0;	Mismatches 1;	Indels 0;
				Gaps 0;

QY	2	MXVAEF	7
D _b	16	MHVAEF	21

RESULT 27
US-09-270-767-32289
; Sequence 32289, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and

```
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 32289
; LENGTH: 285
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-32289

Query Match          79.3%; Score 23; DB 4; Length 285;
Best Local Similarity 57.1%; Pred. No. 4.2e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
   ||:|:|
Db 17 VMQIAQF 23

RESULT 28
US-09-270-767-47506
; Sequence 47506, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 47506
; LENGTH: 285
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-47506

Query Match          79.3%; Score 23; DB 4; Length 285;
Best Local Similarity 57.1%; Pred. No. 4.2e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
   ||:|:|
Db 17 VMQIAQF 23

RESULT 29
US-08-701-191A-13
; Sequence 13, Application US/08701191A
; Patent No. 5942428
; GENERAL INFORMATION:
; APPLICANT: Moosa Mohammadi, Joseph Schlessinger,
; APPLICANT: and Stevan R. Hubbard
; TITLE OF INVENTION: CRYSTALS OF THE TYROSINE KINASE DOMAIN
; TITLE OF INVENTION: OF NON-INSULIN RECEPTOR TYROSINE KINASE
; NUMBER OF SEQUENCES: 41
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FastSeq for Windows 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/701,191A
```

```
; FILING DATE: August 21, 1996
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 227/088
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 299 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-701-191A-13

Query Match          79.3%; Score 23; DB 2; Length 299;
Best Local Similarity 57.1%; Pred. No. 4.5e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
   ||:|:|
Db 172 VMKIADF 178

RESULT 30
US-09-664-526-13
; Sequence 13, Application US/09664526
; Patent No. 6682921
; GENERAL INFORMATION:
; APPLICANT: MOHAMMADI, MOOSA
; APPLICANT: SCHLESSINGER, JOSEPH
; APPLICANT: HUBBARD, STEVAN R.
; TITLE OF INVENTION: CRYSTALS OF THE TYROSINE KINASE DOMAIN OF NON-INSULIN
; TITLE OF INVENTION: RECEPTOR TYROSINE KINASES
; FILE REFERENCE: 038602/0847
; CURRENT APPLICATION NUMBER: US/09/664,526
; CURRENT FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: 09/188,809
; PRIOR FILING DATE: 1998-11-09
; PRIOR APPLICATION NUMBER: 08/701,191
; PRIOR FILING DATE: 1996-08-21
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 299
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-664-526-13

Query Match          79.3%; Score 23; DB 4; Length 299;
Best Local Similarity 57.1%; Pred. No. 4.5e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
   ||:|:|
Db 172 VMKIADF 178

RESULT 31
US-08-701-191A-31
; Sequence 31, Application US/08701191A
; Patent No. 5942428
; GENERAL INFORMATION:
; APPLICANT: Moosa Mohammadi, Joseph Schlessinger,
; APPLICANT: and Stevan R. Hubbard
; TITLE OF INVENTION: CRYSTALS OF THE TYROSINE KINASE DOMAIN
```

; TITLE OF INVENTION: OF NON-INSULIN RECEPTOR TYROSINE KINASE
; NUMBER OF SEQUENCES: 41
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FastSeq for Windows 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/701,191A
; FILING DATE: August 21, 1996
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 227/088
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 300 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-701-191A-31

Query Match 79.3%; Score 23; DB 2; Length 300;
Best Local Similarity 57.1%; Pred. No. 4.5e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VMXVAEF 7
|| :||
Db 172 VKIADF 178

RESULT 32
US-09-664-526-31
; Sequence 31, Application US/09664526
; Patent No. 6682921
; GENERAL INFORMATION:
; APPLICANT: MOHAMMADI, MOOSA
; APPLICANT: SCHLESSINGER, JOSEPH
; APPLICANT: HUBBARD, STEVAN R.
; TITLE OF INVENTION: CRYSTALS OF THE TYROSINE KINASE DOMAIN OF NON-INSULIN
; FILE REFERENCE: 038602/0847
; CURRENT APPLICATION NUMBER: US/09/664,526
; CURRENT FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: 09/188,809
; PRIOR FILING DATE: 1998-11-09
; PRIOR APPLICATION NUMBER: 08/701,191
; PRIOR FILING DATE: 1996-08-21
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 31
; LENGTH: 300
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-664-526-31

Query Match 79.3%; Score 23; DB 4; Length 300;
Best Local Similarity 57.1%; Pred. No. 4.5e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VMXVAEF 7
|| :||
Db 172 VKIADF 178

RESULT 33
US-08-701-191A-9
; Sequence 9, Application US/08701191A
; Patent No. 5942428
; GENERAL INFORMATION:
; APPLICANT: Moosa Mohammadi, Joseph Schlessinger,
; APPLICANT: and Stevan R. Hubbard
; TITLE OF INVENTION: CRYSTALS OF THE TYROSINE KINASE DOMAIN
; TITLE OF INVENTION: OF NON-INSULIN RECEPTOR TYROSINE KINASE
; NUMBER OF SEQUENCES: 41
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FastSeq for Windows 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/701,191A
; FILING DATE: August 21, 1996
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 227/088
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 309 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-701-191A-9

Query Match 79.3%; Score 23; DB 2; Length 309;
Best Local Similarity 57.1%; Pred. No. 4.6e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VMXVAEF 7
|| :||
Db 181 VKIADF 187

RESULT 34
US-09-664-526-9
; Sequence 9, Application US/09664526
; Patent No. 6682921
; GENERAL INFORMATION:
; APPLICANT: MOHAMMADI, MOOSA


```
; APPLICANT: SCHLESSINGER, JOSEPH
; APPLICANT: HUBBARD, STEVAN R.
; TITLE OF INVENTION: CRYSTALS OF THE TYROSINE KINASE DOMAIN OF NON-INSULIN
; TITLE OF INVENTION: RECEPTOR TYROSINE KINASES
; FILE REFERENCE: 038602/0847
; CURRENT APPLICATION NUMBER: US/09/664,526
; CURRENT FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: 09/188,809
; PRIOR FILING DATE: 1998-11-09
; PRIOR APPLICATION NUMBER: 08/701,191
; PRIOR FILING DATE: 1996-08-21
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 309
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-664-526--9

Query Match          79.3%; Score 23; DB 4; Length 309;
Best Local Similarity 57.1%; Pred. No. 4.6e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      1 VMKVAEF 7
      || :||
Db      181 VMKIADF 187

RESULT 35
US-08-701-191A-1
; Sequence 1, Application US/08701191A
; Patent No. 5942428
; GENERAL INFORMATION:
; APPLICANT: Moosa Mohammadi, Joseph Schlessinger,
; APPLICANT: and Stevan R. Hubbard
; TITLE OF INVENTION: CRYSTALS OF THE TYROSINE KINASE DOMAIN
; TITLE OF INVENTION: OF NON-INSULIN RECEPTOR TYROSINE KINASE
; NUMBER OF SEQUENCES: 41
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FastSeq for Windows 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/701,191A
; FILING DATE: August 21, 1996
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 227/088
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 310 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear

; APPLICANT: SCHLESSINGER, JOSEPH
; APPLICANT: HUBBARD, STEVAN R.
; TITLE OF INVENTION: CRYSTALS OF THE TYROSINE KINASE DOMAIN OF NON-INSULIN
; TITLE OF INVENTION: RECEPTOR TYROSINE KINASES
; FILE REFERENCE: 038602/0847
; CURRENT APPLICATION NUMBER: US/09/664,526
; CURRENT FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: 09/188,809
; PRIOR FILING DATE: 1998-11-09
; PRIOR APPLICATION NUMBER: 08/701,191
; PRIOR FILING DATE: 1996-08-21
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 309
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-664-526--9

Query Match          79.3%; Score 23; DB 2; Length 310;
Best Local Similarity 57.1%; Pred. No. 4.6e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      1 VMKVAEF 7
      || :||
Db      181 VMKIADF 187

RESULT 36
US-08-701-191A-6
; Sequence 6, Application US/08701191A
; Patent No. 5942428
; GENERAL INFORMATION:
; APPLICANT: Moosa Mohammadi, Joseph Schlessinger,
; APPLICANT: and Stevan R. Hubbard
; TITLE OF INVENTION: CRYSTALS OF THE TYROSINE KINASE DOMAIN
; TITLE OF INVENTION: OF NON-INSULIN RECEPTOR TYROSINE KINASE
; NUMBER OF SEQUENCES: 41
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FastSeq for Windows 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/701,191A
; FILING DATE: August 21, 1996
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 227/088
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 310 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-701-191A-6

Query Match          79.3%; Score 23; DB 2; Length 310;
Best Local Similarity 57.1%; Pred. No. 4.6e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      1 VMKVAEF 7
      || :||
Db      181 VMKIADF 187

RESULT 37
US-08-701-191A-7
; Sequence 7, Application US/08701191A
```

; Patent No. 5942428
; GENERAL INFORMATION:
; APPLICANT: Moosa Mohammadi, Joseph Schlessinger,
; and Stevan R. Hubbard
; TITLE OF INVENTION: CRYSTALS OF THE TYROSINE KINASE DOMAIN
; OF NON-INSULIN RECEPTOR TYROSINE KINASE
; NUMBER OF SEQUENCES: 41
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FastSeq for Windows 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/701,191A
; FILING DATE: August 21, 1996
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 227/088
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 310 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-701-191A-7

Query Match 79.3%; Score 23; DB 2; Length 310;
Best Local Similarity 57.1%; Pred. No. 4.6e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VMXVAEF 7
Db 181 VMKIADF 187

RESULT 38
US-08-701-191A-8
; Sequence 8, Application US/08701191A
; Patent No. 5942428
; GENERAL INFORMATION:
; APPLICANT: Moosa Mohammadi, Joseph Schlessinger,
; and Stevan R. Hubbard
; TITLE OF INVENTION: CRYSTALS OF THE TYROSINE KINASE DOMAIN
; OF NON-INSULIN RECEPTOR TYROSINE KINASE
; NUMBER OF SEQUENCES: 41
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:

; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FastSeq for Windows 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/701,191A
; FILING DATE: August 21, 1996
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 227/088
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 310 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-701-191A-8

Query Match 79.3%; Score 23; DB 2; Length 310;
Best Local Similarity 57.1%; Pred. No. 4.6e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VMXVAEF 7
Db 181 VMKIADF 187

RESULT 39
US-09-390-326-7
; Sequence 7, Application US/09390326
; Patent No. 6316603
; GENERAL INFORMATION:
; APPLICANT: MCTIGUE, MICHELE A.
; APPLICANT: WICKERSHAM, JOHN A.
; APPLICANT: PINKO, CHRIS
; APPLICANT: SHOWALTER, RICHARD
; APPLICANT: PARAST, CAMRAN V.
; APPLICANT: TEMPCZYK-RUSSEL, ANNA
; APPLICANT: GEHRING, MICHAEL R.
; APPLICANT: MROCKOWSKI, BARBARA
; APPLICANT: KAN, CHEN-CHEN
; APPLICANT: VILLAFRANCA, J. ERNEST
; APPLICANT: APPELT, KRYSZTOF
; TITLE OF INVENTION: MODIFICATIONS OF THE VEGF RECEPTOR-2 PROTEIN AND
; TITLE OF INVENTION: METHODS OF USE
; FILE REFERENCE: 0125-0016US
; CURRENT APPLICATION NUMBER: US/09/390,326
; CURRENT FILING DATE: 1999-09-07
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 310
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-390-326-7

Query Match 79.3%; Score 23; DB 3; Length 310;
Best Local Similarity 57.1%; Pred. No. 4.6e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VMXVAEF 7
||| :||

```
Db      181 VMKIADF 187

RESULT 40
US-09-664-526-1
; Sequence 1, Application US/09664526
; Patent No. 6682921
; GENERAL INFORMATION:
; APPLICANT: MOHAMMADI, MOOSA
; APPLICANT: SCHLESSINGER, JOSEPH
; APPLICANT: HUBBARD, STEVAN R.
; TITLE OF INVENTION: CRYSTALS OF THE TYROSINE KINASE DOMAIN OF NON-INSULIN
; FILE REFERENCE: 038602/0847
; CURRENT APPLICATION NUMBER: US/09/664,526
; CURRENT FILING DATE: 2000-09-18
; PRIOR FILING DATE: 1998-11-09
; PRIOR APPLICATION NUMBER: 08/701,191
; PRIOR FILING DATE: 1996-08-21
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 310
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-664-526-1

Query Match      79.3%; Score 23; DB 4; Length 310;
Best Local Similarity 57.1%; Pred. No. 4.6e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy      1 VMXVAEF 7
Db      181 VMKIADF 187

RESULT 41
US-09-664-526-6
; Sequence 6, Application US/09664526
; Patent No. 6682921
; GENERAL INFORMATION:
; APPLICANT: MOHAMMADI, MOOSA
; APPLICANT: SCHLESSINGER, JOSEPH
; APPLICANT: HUBBARD, STEVAN R.
; TITLE OF INVENTION: CRYSTALS OF THE TYROSINE KINASE DOMAIN OF NON-INSULIN
; FILE REFERENCE: 038602/0847
; CURRENT APPLICATION NUMBER: US/09/664,526
; CURRENT FILING DATE: 2000-09-18
; PRIOR FILING DATE: 1998-11-09
; PRIOR APPLICATION NUMBER: 08/701,191
; PRIOR FILING DATE: 1996-08-21
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 310
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-664-526-6

Query Match      79.3%; Score 23; DB 4; Length 310;
Best Local Similarity 57.1%; Pred. No. 4.6e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy      1 VMXVAEF 7
Db      181 VMKIADF 187

RESULT 42
US-09-664-526-7
; Sequence 7, Application US/09664526
; Patent No. 6682921
; GENERAL INFORMATION:
; APPLICANT: MOHAMMADI, MOOSA
; APPLICANT: SCHLESSINGER, JOSEPH
; APPLICANT: HUBBARD, STEVAN R.
; TITLE OF INVENTION: CRYSTALS OF THE TYROSINE KINASE DOMAIN OF NON-INSULIN
; FILE REFERENCE: 038602/0847
; CURRENT APPLICATION NUMBER: US/09/664,526
; CURRENT FILING DATE: 2000-09-18
; PRIOR FILING DATE: 1998-11-09
; PRIOR APPLICATION NUMBER: 08/701,191
; PRIOR FILING DATE: 1996-08-21
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 310
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-664-526-7

Query Match      79.3%; Score 23; DB 4; Length 310;
Best Local Similarity 57.1%; Pred. No. 4.6e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy      1 VMXVAEF 7
Db      181 VMKIADF 187

RESULT 43
US-09-664-526-8
; Sequence 8, Application US/09664526
; Patent No. 6682921
; GENERAL INFORMATION:
; APPLICANT: MOHAMMADI, MOOSA
; APPLICANT: SCHLESSINGER, JOSEPH
; APPLICANT: HUBBARD, STEVAN R.
; TITLE OF INVENTION: CRYSTALS OF THE TYROSINE KINASE DOMAIN OF NON-INSULIN
; FILE REFERENCE: 038602/0847
; CURRENT APPLICATION NUMBER: US/09/664,526
; CURRENT FILING DATE: 2000-09-18
; PRIOR FILING DATE: 1998-11-09
; PRIOR APPLICATION NUMBER: 08/701,191
; PRIOR FILING DATE: 1996-08-21
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 310
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-664-526-8

Query Match      79.3%; Score 23; DB 4; Length 310;
Best Local Similarity 57.1%; Pred. No. 4.6e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy      1 VMXVAEF 7
Db      181 VMKIADF 187

RESULT 44
US-09-939-833-7
; Sequence 7, Application US/09939833
; Patent No. 6753416
; GENERAL INFORMATION:
; APPLICANT: MCTIGUE, MICHELE A.
; APPLICANT: WICKERSHAM, JOHN A.
```

```
; Sequence 7, Application US/09664526
; Patent No. 6682921
; GENERAL INFORMATION:
; APPLICANT: MOHAMMADI, MOOSA
; APPLICANT: SCHLESSINGER, JOSEPH
; APPLICANT: HUBBARD, STEVAN R.
; TITLE OF INVENTION: CRYSTALS OF THE TYROSINE KINASE DOMAIN OF NON-INSULIN
; FILE REFERENCE: 038602/0847
; CURRENT APPLICATION NUMBER: US/09/664,526
; CURRENT FILING DATE: 2000-09-18
; PRIOR FILING DATE: 1998-11-09
; PRIOR APPLICATION NUMBER: 08/701,191
; PRIOR FILING DATE: 1996-08-21
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 310
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-664-526-7

Query Match      79.3%; Score 23; DB 4; Length 310;
Best Local Similarity 57.1%; Pred. No. 4.6e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy      1 VMXVAEF 7
Db      181 VMKIADF 187

RESULT 43
US-09-664-526-8
; Sequence 8, Application US/09664526
; Patent No. 6682921
; GENERAL INFORMATION:
; APPLICANT: MOHAMMADI, MOOSA
; APPLICANT: SCHLESSINGER, JOSEPH
; APPLICANT: HUBBARD, STEVAN R.
; TITLE OF INVENTION: CRYSTALS OF THE TYROSINE KINASE DOMAIN OF NON-INSULIN
; FILE REFERENCE: 038602/0847
; CURRENT APPLICATION NUMBER: US/09/664,526
; CURRENT FILING DATE: 2000-09-18
; PRIOR FILING DATE: 1998-11-09
; PRIOR APPLICATION NUMBER: 08/701,191
; PRIOR FILING DATE: 1996-08-21
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 310
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-664-526-8

Query Match      79.3%; Score 23; DB 4; Length 310;
Best Local Similarity 57.1%; Pred. No. 4.6e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy      1 VMXVAEF 7
Db      181 VMKIADF 187

RESULT 44
US-09-939-833-7
; Sequence 7, Application US/09939833
; Patent No. 6753416
; GENERAL INFORMATION:
; APPLICANT: MCTIGUE, MICHELE A.
; APPLICANT: WICKERSHAM, JOHN A.
```

```

; APPLICANT: PINKO, CHRIS
; APPLICANT: SHOWALTER, RICHARD
; APPLICANT: PARAST, CAMRAN V.
; APPLICANT: TEMPZYK-RUSSEL, ANNA
; APPLICANT: GEHRING, MICHAEL R.
; APPLICANT: MROCZKOWSKI, BARBARA
; APPLICANT: KAN, CHEN-CHEN
; APPLICANT: VILLAFRANCA, J. ERNEST
; APPLICANT: APPELT, KRZYSZTOF
; TITLE OF INVENTION: MODIFICATIONS OF THE VEGF RECEPTOR-2 PROTEIN AND
; TITLE OF INVENTION: METHODS OF USE
; FILE REFERENCE: 0125-0016US
; CURRENT APPLICATION NUMBER: US/09/939,833
; CURRENT FILING DATE: 2001-08-28
; PRIOR APPLICATION NUMBER: US/09/390,326
; PRIOR FILING DATE: 1999-09-07
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 310
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-939-833-7

Query Match          79.3%; Score 23; DB 4; Length 310;
Best Local Similarity 57.1%; Pred. No. 4.6e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      1 VMKVAEF 7
Db      181 VMKIADP 187

RESULT 45
US-09-506-906-7
; Sequence 7, Application US/09506906
; Patent No. 6784285
; GENERAL INFORMATION:
; APPLICANT: MCTIGUE, MICHELE A.
; APPLICANT: WICKERSHAM, JOHN A.
; APPLICANT: PINKO, CHRIS
; APPLICANT: SHOWALTER, RICHARD
; APPLICANT: PARAST, CAMRAN V.
; APPLICANT: TEMPZYK-RUSSEL, ANNA
; APPLICANT: GEHRING, MICHAEL R.
; APPLICANT: MROCZKOWSKI, BARBARA
; APPLICANT: KAN, CHEN-CHEN
; APPLICANT: VILLAFRANCA, J. ERNEST
; APPLICANT: APPELT, KRZYSZTOF
; TITLE OF INVENTION: MODIFICATIONS OF THE VEGF RECEPTOR-2 PROTEIN AND
; TITLE OF INVENTION: METHODS OF USE
; FILE REFERENCE: 0125-0016US
; CURRENT APPLICATION NUMBER: US/09/506,906
; CURRENT FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: US/09/390,326
; PRIOR FILING DATE: 1999-09-07
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 310
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-506-906-7

Query Match          79.3%; Score 23; DB 4; Length 310;
Best Local Similarity 57.1%; Pred. No. 4.6e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      1 VMKVAEF 7
Db      181 VMKIADP 187

; APPLICANT: PINKO, CHRIS
; APPLICANT: SHOWALTER, RICHARD
; APPLICANT: PARAST, CAMRAN V.
; APPLICANT: TEMPZYK-RUSSEL, ANNA
; APPLICANT: GEHRING, MICHAEL R.
; APPLICANT: MROCZKOWSKI, BARBARA
; APPLICANT: KAN, CHEN-CHEN
; APPLICANT: VILLAFRANCA, J. ERNEST
; APPLICANT: APPELT, KRZYSZTOF
; TITLE OF INVENTION: MODIFICATIONS OF THE VEGF RECEPTOR-2 PROTEIN AND
; TITLE OF INVENTION: METHODS OF USE
; FILE REFERENCE: 0125-0016US
; CURRENT APPLICATION NUMBER: US/09/506,906
; CURRENT FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: US/09/390,326
; PRIOR FILING DATE: 1999-09-07
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 310
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-506-906-7

Query Match          79.3%; Score 23; DB 4; Length 310;
Best Local Similarity 57.1%; Pred. No. 4.6e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      1 VMKVAEF 7
Db      181 VMKIADP 187
```

```

RESULT 46
US-09-939-832-7
; Sequence 7, Application US/09939832
; Patent No. 6794146
; GENERAL INFORMATION:
; APPLICANT: MCTIGUE, MICHELE A.
; APPLICANT: WICKERSHAM, JOHN A.
; APPLICANT: PINKO, CHRIS
; APPLICANT: SHOWALTER, RICHARD
; APPLICANT: PARAST, CAMRAN V.
; APPLICANT: TEMPZYK-RUSSEL, ANNA
; APPLICANT: GEHRING, MICHAEL R.
; APPLICANT: MROCZKOWSKI, BARBARA
; APPLICANT: KAN, CHEN-CHEN
; APPLICANT: VILLAFRANCA, J. ERNEST
; APPLICANT: APPELT, KRZYSZTOF
; TITLE OF INVENTION: MODIFICATIONS OF THE VEGF RECEPTOR-2 PROTEIN AND
; TITLE OF INVENTION: METHODS OF USE
; FILE REFERENCE: 0125-0016US
; CURRENT APPLICATION NUMBER: US/09/939,832
; CURRENT FILING DATE: 2001-08-28
; PRIOR APPLICATION NUMBER: US/09/390,326
; PRIOR FILING DATE: 1999-09-07
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 310
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-939-832-7

Query Match          79.3%; Score 23; DB 4; Length 310;
Best Local Similarity 57.1%; Pred. No. 4.6e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      1 VMKVAEF 7
Db      181 VMKIADP 187

RESULT 47
US-08-278-089A-17
; Sequence 17, Application US/08278089A
; Patent No. 5681714
; GENERAL INFORMATION:
; APPLICANT: BREITMAN, MARTIN L.
; APPLICANT: ROSSANT, JANET
; APPLICANT: DUMONT, DANIEL J.
; APPLICANT: YAMAGUCHI, TERRY P.
; TITLE OF INVENTION: No. 5681714e1 Receptor Tyrosine Kinase
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bereskin & Parr
; STREET: 40 King Street West
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5H 3Y2
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/278,089A
; FILING DATE: 20-JUL-1994
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Kurdzydk, Linda M.
; REGISTRATION NUMBER: 34,971
; REFERENCE/DOCKET NUMBER: 3153-111
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 364-7311
```

TELEFAX: (416) 361-1398
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 313 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; IMMEDIATE SOURCE:
; CLONE: Flgm
US-08-278-089A-17

Query Match 79.3%; Score 23; DB 1; Length 313;
Best Local Similarity 57.1%; Pred. No. 4.7e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VMXVAEF 7
||:|:
Db 159 VKMIADF 165

RESULT 48
US-08-838-957A-16
; Sequence 16, Application US/08838957A
; Patent No. 5998187
; GENERAL INFORMATION:
; APPLICANT: Breitman, Martin L.
; APPLICANT: Rossant, Janet
; APPLICANT: Dumont, Daniel J.
; APPLICANT: Yamaguchi, Terry P.
; TITLE OF INVENTION: No. 5998187el Receptor Tyrosine Kinase
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bereskin & Parr
; STREET: 40 King Street West
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5H 3Y2
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/838,957A
; FILING DATE: 23-APR-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Kurdydk, Linda M.
; REGISTRATION NUMBER: 34,971
; REFERENCE/DOCKET NUMBER: 3153-212
; TELEPHONE: (416) 364-7311
; TELEFAX: (416) 361-1398
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 313 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; IMMEDIATE SOURCE:
; CLONE: Flgm
US-08-838-957A-16

Query Match 79.3%; Score 23; DB 2; Length 313;
Best Local Similarity 57.1%; Pred. No. 4.7e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VMXVAEF 7
||:|:
Db 159 VKMIADF 165

RESULT 49
US-08-701-191A-2
; Sequence 2, Application US/08701191A
; Patent No. 5942428
; GENERAL INFORMATION:
; APPLICANT: Moosa Mohammadi, Joseph Schlessinger,
; APPLICANT: and Stevan R. Hubbard
; TITLE OF INVENTION: CRYSTALS OF THE TYROSINE KINASE DOMAIN
; TITLE OF INVENTION: OF NON-INSULIN RECEPTOR TYROSINE KINASE
; NUMBER OF SEQUENCES: 41
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: Storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FastSeq for Windows 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/701,191A
; FILING DATE: August 21, 1996
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 227/088
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 315 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
US-08-701-191A-2

Query Match 79.3%; Score 23; DB 2; Length 315;
Best Local Similarity 57.1%; Pred. No. 4.7e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VMXVAEF 7
||:|:
Db 186 VKMIADF 192

RESULT 50
US-09-664-526-2
; Sequence 2, Application US/09664526
; Patent No. 6682921
; GENERAL INFORMATION:
; APPLICANT: MOHAMMADI, MOOSA
; APPLICANT: SCHLESSINGER, JOSEPH
; APPLICANT: HUBBARD, STEVAN R.
; TITLE OF INVENTION: CRYSTALS OF THE TYROSINE KINASE DOMAIN OF NON-INSULIN
; TITLE OF INVENTION: RECEPTOR TYROSINE KINASES
; FILE REFERENCE: 038602/0847
; CURRENT APPLICATION NUMBER: US/09/664,526
; CURRENT FILING DATE: 2000-09-18

;; PRIOR APPLICATION NUMBER: 09/188,809
;; PRIOR FILING DATE: 1998-11-09
;; PRIOR APPLICATION NUMBER: 08/701,191
;; PRIOR FILING DATE: 1996-08-21
;; NUMBER OF SEQ ID NOS: 41
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 2
;; LENGTH: 315
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-664-526-2

Query Match 79.3%; Score 23; DB 4; Length 315;
Best Local Similarity 57.1%; Pred. No. 4.7e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMKVAEF 7
||:|:
Db 186 VMKIADF 192

RESULT 51
US-08-701-191A-10
; Sequence 10, Application US/08701191A
; Patent No. 5942428
; GENERAL INFORMATION:
; APPLICANT: Moosa Mohammadi, Joseph Schlessinger,
; APPLICANT: and Stevan R. Hubbard
; TITLE OF INVENTION: CRYSTALS OF THE TYROSINE KINASE DOMAIN
; TITLE OF INVENTION: OF NON-INSULIN RECEPTOR TYROSINE KINASE
; NUMBER OF SEQUENCES: 41
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; CITY: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FastSeq for Windows 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/701,191A
; FILING DATE: August 21, 1996
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 227/088
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 318 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-701-191A-10

Query Match 79.3%; Score 23; DB 2; Length 318;
Best Local Similarity 57.1%; Pred. No. 4.8e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMKVAEF 7
||:|:
Db 188 VMKIADF 194

RESULT 52
US-09-664-526-10
; Sequence 10, Application US/09664526
; Patent No. 6682921
; GENERAL INFORMATION:
; APPLICANT: MOHAMMADI, MOOSA
; APPLICANT: SCHLESSINGER, JOSEPH
; APPLICANT: HUBBARD, STEVAN R.
; TITLE OF INVENTION: CRYSTALS OF THE TYROSINE KINASE DOMAIN OF NON-INSULIN
; TITLE OF INVENTION: RECEPTOR TYROSINE KINASES
; FILE REFERENCE: 038602/0847
; CURRENT APPLICATION NUMBER: US/09/664,526
; CURRENT FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: 09/188,809
; PRIOR FILING DATE: 1998-11-09
; PRIOR APPLICATION NUMBER: 08/701,191
; PRIOR FILING DATE: 1996-08-21
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 318
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-664-526-10

Query Match 79.3%; Score 23; DB 4; Length 318;
Best Local Similarity 57.1%; Pred. No. 4.8e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMKVAEF 7
||:|:
Db 188 VMKIADF 194

RESULT 53
US-09-134-001C-5567
; Sequence 5567, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 5567
; LENGTH: 344
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-5567

Query Match 79.3%; Score 23; DB 3; Length 344;
Best Local Similarity 57.1%; Pred. No. 5.2e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMKVAEF 7
||:|:
Db 303 VMIIADF 309

RESULT 54
US-08-701-191A-3
; Sequence 3, Application US/08701191A

; Patent No. 5942428
; GENERAL INFORMATION:
; APPLICANT: Moosa Mohammadi, Joseph Schlessinger,
; APPLICANT: and Stevan R. Hubbard
; TITLE OF INVENTION: CRYSTALS OF THE TYROSINE KINASE DOMAIN
; TITLE OF INVENTION: OF NON-INSULIN RECEPTOR TYROSINE KINASE
; NUMBER OF SEQUENCES: 41
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FastSEQ for Windows 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/701,191A
; FILING DATE: August 21, 1996
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 227/088
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 351 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; US-08-701-191A-3

Query Match 79.3%; Score 23; DB 2; Length 351;
Best Local Similarity 57.1%; Pred. No. 5.3e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
||| :||
Db 222 VKIADF 228

RESULT 55
US-09-664-526-3
; Sequence 3, Application US/09664526
; Patent No. 6682921
; GENERAL INFORMATION:
; APPLICANT: MOHAMMADI, MOOSA
; APPLICANT: SCHLESSINGER, JOSEPH
; APPLICANT: HUBBARD, STEVAN R.
; TITLE OF INVENTION: CRYSTALS OF THE TYROSINE KINASE DOMAIN OF NON-INSULIN
; TITLE OF INVENTION: RECEPTOR TYROSINE KINASES
; FILE REFERENCE: 038602/0847
; CURRENT APPLICATION NUMBER: US/09/664,526
; CURRENT FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: 09/188,809
; PRIOR FILING DATE: 1998-11-09
; PRIOR APPLICATION NUMBER: 08/701,191
; PRIOR FILING DATE: 1996-08-21
; NUMBER OF SEQ ID NOS: 41

; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: protein encoded by recombinant baculovirus
US-09-664-526-3

Query Match 79.3%; Score 23; DB 4; Length 351;
Best Local Similarity 57.1%; Pred. No. 5.3e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
||| :||
Db 222 VKIADF 228

RESULT 56
US-08-070-165F-8
; Sequence 8, Application US/08070165F
; Patent No. 5750365
; GENERAL INFORMATION:
; APPLICANT: Chiu, Ing-Ming
; APPLICANT: Poulin, Matthew L
; TITLE OF INVENTION: Acidic Fibroblast Growth Factor (aFGF)
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ing-Ming Chiu
; STREET: S2052 Davis Medical Research Center, 480 West
; STREET: 9th Avenue
; CITY: Columbus
; STATE: Ohio
; COUNTRY: USA
; ZIP: 43210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/070,165F
; FILING DATE:
; CLASSIFICATION: 435
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (614)-293-8093
; TELEFAX: (614)-293-5631
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 378 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-070-165F-8

Query Match 79.3%; Score 23; DB 1; Length 378;
Best Local Similarity 57.1%; Pred. No. 5.7e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
||| :||
Db 197 VKIADF 203

RESULT 57
US-08-885-418-8
; Sequence 8, Application US/08885418
; Patent No. 5925528
; GENERAL INFORMATION:
; APPLICANT: Chiu, Ing-Ming
; APPLICANT: Poulin, Matthew L
; TITLE OF INVENTION: Acidic Fibroblast Growth Factor (aFGF)

```
;
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ing-Ming Chiu
; STREET: S2052 Davis Medical Research Center, 480 West
; STREET: 9th Avenue
; CITY: Columbus
; STATE: Ohio
; COUNTRY: USA
; ZIP: 43210
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/885,418
; FILING DATE:
; CLASSIFICATION: 435
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (614)-293-8093
; TELEFAX: (614)-293-5631
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 378 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
US-08-885-418-8
;
Query Match 79.3%; Score 23; DB 2; Length 378;
Best Local Similarity 57.1%; Pred. No. 5.7e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VMXVAEF 7
Db 197 VMKIADF 203

RESULT 58
US-08-070-165F-4
; Sequence 4, Application US/08070165F
; Patent No. 5750365
; GENERAL INFORMATION:
; APPLICANT: Chiu, Ing-Ming L
; APPLICANT: Poulin, Matthew L
; TITLE OF INVENTION: Acidic Fibroblast Growth Factor (aFGF)
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ing-Ming Chiu
; STREET: S2052 Davis Medical Research Center, 480 West
; STREET: 9th Avenue
; CITY: Columbus
; STATE: Ohio
; COUNTRY: USA
; ZIP: 43210
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/070,165F
; FILING DATE:
; CLASSIFICATION: 435
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (614)-293-8093
; TELEFAX: (614)-293-5631
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 388 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
```

```
US-08-070-165F-4
;
Query Match 79.3%; Score 23; DB 1; Length 388;
Best Local Similarity 57.1%; Pred. No. 5.9e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VMXVAEF 7
Db 204 VMKIADF 210

RESULT 59
US-08-885-418-4
; Sequence 4, Application US/08885418
; Patent No. 5925528
; GENERAL INFORMATION:
; APPLICANT: Chiu, Ing-Ming L
; APPLICANT: Poulin, Matthew L
; TITLE OF INVENTION: Acidic Fibroblast Growth Factor (aFGF)
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ing-Ming Chiu
; STREET: S2052 Davis Medical Research Center, 480 West
; STREET: 9th Avenue
; CITY: Columbus
; STATE: Ohio
; COUNTRY: USA
; ZIP: 43210
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/885,418
; FILING DATE:
; CLASSIFICATION: 435
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (614)-293-8093
; TELEFAX: (614)-293-5631
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 388 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
US-08-885-418-4
;
Query Match 79.3%; Score 23; DB 2; Length 388;
Best Local Similarity 57.1%; Pred. No. 5.9e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VMXVAEF 7
Db 204 VMKIADF 210

RESULT 60
US-09-252-991A-28422
; Sequence 28422, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
```



```
; SEQ ID NO 28422
; LENGTH: 397
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-28422

Query Match      79.3%; Score 23; DB 4; Length 397;
Best Local Similarity 66.7%; Pred. No. 6e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      2 MXVAEF 7
Db      259 MAIAEF 264

RESULT 61
US-09-270-767-43247
; Sequence 43247, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 43247
; LENGTH: 412
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-43247

Query Match      79.3%; Score 23; DB 4; Length 412;
Best Local Similarity 71.4%; Pred. No. 6.3e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 VMXVAEF 7
Db      176 VMLVVEF 182

RESULT 62
US-09-710-279-1818
; Sequence 1818, Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: P03480US
; CURRENT APPLICATION NUMBER: US/09/710,279
; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1818
; LENGTH: 421
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
US-09-710-279-1818

Query Match      79.3%; Score 23; DB 4; Length 421;
Best Local Similarity 83.3%; Pred. No. 6.4e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 MXVAEF 7
Db      176 VMLVVEF 182

RESULT 63
US-10-138-701-8
; Sequence 8, Application US/10138701
; Patent No. 6753149
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc. et al.
; TITLE OF INVENTION: Staphylococcus aureus genes and polypeptides
; FILE REFERENCE: PB484
; CURRENT APPLICATION NUMBER: US/10/138,701
; CURRENT FILING DATE: 2002-05-06
; PRIOR APPLICATION NUMBER: US/09/512,255A
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/098,964
; PRIOR FILING DATE: 1998-09-01
; PRIOR APPLICATION NUMBER: US 60/009,861
; PRIOR FILING DATE: 1996-01-05
; PRIOR APPLICATION NUMBER: PCT/ US99/19726
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: US 08/956,171
; PRIOR FILING DATE: 1997-10-20
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 8
; LENGTH: 421
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-10-138-701-8

Query Match      79.3%; Score 23; DB 4; Length 421;
Best Local Similarity 83.3%; Pred. No. 6.4e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 MXVAEF 7
Db      336 MHVAEF 341

RESULT 64
US-10-138-701-22
; Sequence 22, Application US/10138701
; Patent No. 6753149
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc. et al.
; TITLE OF INVENTION: Staphylococcus aureus genes and polypeptides
; FILE REFERENCE: PB484
; CURRENT APPLICATION NUMBER: US/10/138,701
; CURRENT FILING DATE: 2002-05-06
; PRIOR APPLICATION NUMBER: US/09/512,255A
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/098,964
; PRIOR FILING DATE: 1998-09-01
; PRIOR APPLICATION NUMBER: US 60/009,861
; PRIOR FILING DATE: 1996-01-05
; PRIOR APPLICATION NUMBER: PCT/ US99/19726
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: US 08/956,171
; PRIOR FILING DATE: 1997-10-20
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 22
; LENGTH: 421
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-10-138-701-22

Query Match      79.3%; Score 23; DB 4; Length 421;
Best Local Similarity 83.3%; Pred. No. 6.4e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 MXVAEF 7
Db      336 MHVAEF 341
```

```
Db          336 MHVAF 341
| |||
RESULT 65
US-09-134-001C-4257
; Sequence 4257, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 4257
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4257
Query Match          79.3%; Score 23; DB 3; Length 431;
Best Local Similarity 83.3%; Pred. No. 6.6e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy          2 MXVAF 7
| |||
Db          346 MHVAF 351
| |||
RESULT 66
US-09-252-991A-31250
; Sequence 31250, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 31250
; LENGTH: 483
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-31250
Query Match          79.3%; Score 23; DB 4; Length 483;
Best Local Similarity 57.1%; Pred. No. 7.4e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy          1 VMXVAF 7
| : |||
Db          433 VLGIAF 439
| : |||
RESULT 67
US-08-673-789-11
; Sequence 11, Application US/08673789
; Patent No. 5814479
; GENERAL INFORMATION:
; APPLICANT: ZHOU, RENPING; SCHULZ, NICHOLAS,
; APPLICANT: T.; KROMER, LAWRENCE, F.; VANDE WOUDE,
```

```
; APPLICANT: GEORGE, F.
; TITLE OF INVENTION: BSK RECEPTOR LIKE
; TITLE OF INVENTION: TYROSINE KINASE AND LIGAND AND THEIR
; TITLE OF INVENTION: USE IN DIAGNOSTIC AND THERAPEUTIC
; TITLE OF INVENTION: METHODS
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSER: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/673,789
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/177,812
; FILING DATE: 04-JAN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: CAROL M. GRUPPI
; REGISTRATION NUMBER: 37,341
; REFERENCE/DOCKET NUMBER: 2026-4105
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 612
; TYPE: AMINO ACID
; STRANDEDNESS: UNKNOWN
; TOPOLOGY: UNKNOWN
; US-08-673-789-11
Query Match          79.3%; Score 23; DB 2; Length 612;
Best Local Similarity 71.4%; Pred. No. 9.6e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy          1 VMXVAF 7
| |||
Db          321 VMIVTEF 327
| |||
RESULT 68
US-09-328-352-7991
; Sequence 7991, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 7991
; LENGTH: 613
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-7991
Query Match          79.3%; Score 23; DB 4; Length 613;
Best Local Similarity 57.1%; Pred. No. 9.6e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy          1 VMXVAF 7
```

```
Db          :| :|||
53 IMELAEF 59

RESULT 69
US-08-471-570-10
; Sequence 10, Application US/08471570
; Patent No. 5750371
; GENERAL INFORMATION:
; APPLICANT: IGARASHI, Koichi
; APPLICANT: SENO, Masaharu
; APPLICANT: WATANABE, Tatsuya
; TITLE OF INVENTION: PROTEIN, DNA AND USE THEREOF
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DAVID G. CONLIN, DIKE, BRONSTEIN, ROBERTS &
; ADDRESSEE: CUSHMAN
; STREET: 130 Water Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: US
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/471,570
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/149,664
; FILING DATE:
; APPLICATION NUMBER: US 07/743369
; FILING DATE: 16-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: LINEK, Ernest V
; REGISTRATION NUMBER: 29822
; REFERENCE/DOCKET NUMBER: 40897
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)523-3400
; TELEFAX: (617)523-6440
; TELEX: 200291 STRE UR
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 652 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-471-570-10

Query Match          79.3%; Score 23; DB 1; Length 652;
Best Local Similarity 57.1%; Pred. No. 1e+03;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy          1 VMXVAEF 7
Db          523 VMKIADF 529

RESULT 70
US-07-640-029-3
; Sequence 3, Application US/07640029
; Patent No. 5229501
; GENERAL INFORMATION:
; APPLICANT: Kiefer, Michael C.
; APPLICANT: Valenzuela, Pablo D.T.
; APPLICANT: Bart, Philip J.
; TITLE OF INVENTION: Expression and Use of Human Fibroblast
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:

US-08-471-570-10
; Sequence 10, Application US/08471570
; Patent No. 5750371
; GENERAL INFORMATION:
; APPLICANT: IGARASHI, Koichi
; APPLICANT: SENO, Masaharu
; APPLICANT: WATANABE, Tatsuya
; TITLE OF INVENTION: PROTEIN, DNA AND USE THEREOF
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DAVID G. CONLIN, DIKE, BRONSTEIN, ROBERTS &
; ADDRESSEE: CUSHMAN
; STREET: 130 Water Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: US
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/471,570
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/149,664
; FILING DATE:
; APPLICATION NUMBER: US 07/743369
; FILING DATE: 16-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: LINEK, Ernest V
; REGISTRATION NUMBER: 29822
; REFERENCE/DOCKET NUMBER: 40897
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)523-3400
; TELEFAX: (617)523-6440
; TELEX: 200291 STRE UR
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 652 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-471-570-10

Query Match          79.3%; Score 23; DB 1; Length 652;
Best Local Similarity 57.1%; Pred. No. 1e+03;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy          1 VMXVAEF 7
Db          523 VMKIADF 529

RESULT 71
US-08-070-165F-6
; Sequence 6, Application US/08070165F
; Patent No. 5750365
; GENERAL INFORMATION:
; APPLICANT: Chiu, Ing-Ming
; APPLICANT: Poulin, Matthew L.
; TITLE OF INVENTION: Acidic Fibroblast Growth Factor (aFGF)
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ing-Ming Chiu
; STREET: 52052 Davis Medical Research Center, 480 West
; STREET: 9th Avenue
; CITY: Columbus
; STATE: Ohio
; COUNTRY: USA
; ZIP: 43210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/070,165F
; FILING DATE:
; CLASSIFICATION: 435
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (614)-293-8093
; TELEFAX: (614)-293-5631
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 729 amino acids
; TYPE: amino acid
US-07-640-029-3

Query Match          79.3%; Score 23; DB 1; Length 729;
Best Local Similarity 57.1%; Pred. No. 1.2e+03;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy          1 VMXVAEF 7
Db          543 VMKIADF 549

RESULT 71
US-08-070-165F-6
; Sequence 6, Application US/08070165F
; Patent No. 5750365
; GENERAL INFORMATION:
; APPLICANT: Chiu, Ing-Ming
; APPLICANT: Poulin, Matthew L.
; TITLE OF INVENTION: Acidic Fibroblast Growth Factor (aFGF)
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ing-Ming Chiu
; STREET: 52052 Davis Medical Research Center, 480 West
; STREET: 9th Avenue
; CITY: Columbus
; STATE: Ohio
; COUNTRY: USA
; ZIP: 43210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/070,165F
; FILING DATE:
; CLASSIFICATION: 435
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (614)-293-8093
; TELEFAX: (614)-293-5631
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 729 amino acids
; TYPE: amino acid
US-07-640-029-3
```

```
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-070-165F-6

Query Match      79.3%; Score 23; DB 1; Length 729;
Best Local Similarity 57.1%; Pred. No. 1.2e+03;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy      1 VMXVAEF 7
Db      547 VMKIADF 553

RESULT 72
US-08-885-418-6
; Sequence 6, Application US/08885418
; Patent No. 5925528
; GENERAL INFORMATION:
; APPLICANT: Chiu, Ing-Ming L
; APPLICANT: Poulin, Matthew L
; TITLE OF INVENTION: Acidic Fibroblast Growth Factor (aFGF)
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ing-Ming Chiu
; STREET: S2052 Davis Medical Research Center, 480 West
; STREET: 9th Avenue
; CITY: Columbus
; STATE: Ohio
; COUNTRY: USA
; ZIP: 43210
; COMPUTER READABLE FORM:
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: US/08/885,418
; CLASSIFICATION: 435
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (614)-293-8093
; TELEFAX: (614)-293-5631
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 729 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-885-418-6

Query Match      79.3%; Score 23; DB 2; Length 729;
Best Local Similarity 57.1%; Pred. No. 1.2e+03;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy      1 VMXVAEF 7
Db      547 VMKIADF 553

RESULT 73
US-07-921-807B-5
; Sequence 5, Application US/07921807B
; Patent No. 5474914
; GENERAL INFORMATION:
; APPLICANT: SPAETE, RICHARD
; APPLICANT: SPAETE, RICHARD
; TITLE OF INVENTION: METHOD OF INCREASING EXPRESSION
; TITLE OF INVENTION: OF VIRAL PROTEINS
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CHIRON CORPORATION
; STREET: 4560 Horton Street - R440
; CITY: Emeryville
; STATE: CA

; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-070-165F-6

Query Match      79.3%; Score 23; DB 1; Length 729;
Best Local Similarity 57.1%; Pred. No. 1.2e+03;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy      1 VMXVAEF 7
Db      545 VMKIADF 551

RESULT 74
US-08-070-165F-10
; Sequence 10, Application US/08070165F
; Patent No. 5750365
; GENERAL INFORMATION:
; APPLICANT: Chiu, Ing-Ming L
; APPLICANT: Poulin, Matthew L
; TITLE OF INVENTION: Acidic Fibroblast Growth Factor (aFGF)
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ing-Ming Chiu
; STREET: S2052 Davis Medical Research Center, 480 West
; STREET: 9th Avenue
; CITY: Columbus
; STATE: Ohio
; COUNTRY: USA
; ZIP: 43210
; COMPUTER READABLE FORM:
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: US/08/070,165F
; CLASSIFICATION: 435
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (614)-293-8093
; TELEFAX: (614)-293-5631
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 731 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-070-165F-10

Query Match      79.3%; Score 23; DB 1; Length 731;
Best Local Similarity 57.1%; Pred. No. 1.2e+03;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy      1 VMXVAEF 7
Db      545 VMKIADF 551

RESULT 75
US-08-070-165F-10
; Sequence 10, Application US/08070165F
; Patent No. 5750365
; GENERAL INFORMATION:
; APPLICANT: Chiu, Ing-Ming L
; APPLICANT: Poulin, Matthew L
; TITLE OF INVENTION: Acidic Fibroblast Growth Factor (aFGF)
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ing-Ming Chiu
; STREET: S2052 Davis Medical Research Center, 480 West
; STREET: 9th Avenue
; CITY: Columbus
; STATE: Ohio
; COUNTRY: USA
; ZIP: 43210
; COMPUTER READABLE FORM:
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: US/08/070,165F
; CLASSIFICATION: 435
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (614)-293-8093
; TELEFAX: (614)-293-5631
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 731 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-070-165F-10
```

Query Match 79.3%; Score 23; DB 1; Length 731;
Best Local Similarity 57.1%; Pred. No. 1.2e+03;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEP 7
DB 549 VMKIADF 555

RESULT 75

US-08-441-944A-5
; Sequence 5, Application US/08441944A
; Patent No. 5767250
; GENERAL INFORMATION:
; APPLICANT: SPAETE, RICHARD
; TITLE OF INVENTION: METHOD OF INCREASING EXPRESSION
; TITLE OF INVENTION: OF VIRAL PROTEINS
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CHIRON CORPORATION
; STREET: 4560 Horton Street - R440
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608-2916
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/441,944A
; FILING DATE: 16-MAY-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/921,807
; FILING DATE: 29-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: MCCLUNG, BARBARA G.
; REGISTRATION NUMBER: 33,113
; REFERENCE/DOCKET NUMBER: 0209.001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 601-2708
; TELEFAX: (510) 655-3542
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 731 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-441-944A-5

Query Match 79.3%; Score 23; DB 1; Length 731;
Best Local Similarity 57.1%; Pred. No. 1.2e+03;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEP 7
DB 545 VMKIADF 551

RESULT 76

US-08-885-418-10
; Sequence 10, Application US/08885418
; Patent No. 5925528
; GENERAL INFORMATION:
; APPLICANT: Chiu, Ing-Ming
; APPLICANT: Poulin, Matthew L
; TITLE OF INVENTION: Acidic Fibroblast Growth Factor (aFGF)
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ing-Ming Chiu

STREET: S2052 Davis Medical Research Center, 480 West
STREET: 9th Avenue
CITY: Columbus
STATE: Ohio
COUNTRY: USA
ZIP: 43210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/885,418
; FILING DATE:
; CLASSIFICATION: 435
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (614)-293-8093
; TELEFAX: (614)-293-5631
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 731 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-885-418-10

Query Match 79.3%; Score 23; DB 2; Length 731;
Best Local Similarity 57.1%; Pred. No. 1.2e+03;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEP 7
DB 549 VMKIADF 555

RESULT 77

US-08-439-992A-3
; Sequence 3, Application US/08439992A
; Patent No. 6255454
; GENERAL INFORMATION:
; APPLICANT: Kiefer, Michael C.
; APPLICANT: Pablo, Valenzuela D.T.
; APPLICANT: Philip, Barr J.
; TITLE OF INVENTION: Expression and Use of Human Fibroblast
; TITLE OF INVENTION: Receptor
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/439,992A
; FILING DATE: 12-MAY-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Chung, Ling-Pong
; REGISTRATION NUMBER: 36,482
; REFERENCE/DOCKET NUMBER: 0165.004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 510-923-2704
; TELEFAX: (510) 655-3542
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 731 amino acids
; TYPE: amino acid

Query Match 79.3%; Score 23; DB 1; Length 731;
Best Local Similarity 57.1%; Pred. No. 1.2e+03;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEP 7
DB 545 VMKIADF 551

RESULT 76

US-08-885-418-10
; Sequence 10, Application US/08885418
; Patent No. 5925528
; GENERAL INFORMATION:
; APPLICANT: Chiu, Ing-Ming
; APPLICANT: Poulin, Matthew L
; TITLE OF INVENTION: Acidic Fibroblast Growth Factor (aFGF)
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ing-Ming Chiu

```
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-439-992A-3

Query Match          79.3%; Score 23; DB 3; Length 731;
Best Local Similarity 57.1%; Pred. No. 1.2e+03;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      1 VMXVAEF 7
Db      545 VMKIADF 551

RESULT 78
US-07-640-029--4
; Sequence 4, Application US/07640029
; Patent No. 5229501
; GENERAL INFORMATION:
; APPLICANT: Kiefer, Michael C.
; APPLICANT: Valenzuela, Pablo D.T.
; APPLICANT: Barr, Philip J.
; TITLE OF INVENTION: Expression and Use of Human Fibroblast
; TITLE OF INVENTION: Growth Factor Receptor
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: California
; COUNTRY: USA
; ZIP: 94608
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/640,029
; FILING DATE: 19910111
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: McClung, Barbara G.
; REGISTRATION NUMBER: 33,113
; REFERENCE/DOCKET NUMBER: CH-165
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 510-601-2708
; TELEFAX: 510-655-3542
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 733 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-07-640-029--4

Query Match          79.3%; Score 23; DB 1; Length 733;
Best Local Similarity 57.1%; Pred. No. 1.2e+03;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      1 VMXVAEF 7
Db      547 VMKIADF 553

RESULT 79
US-07-921-807B-6
; Sequence 6, Application US/07921807B
; Patent No. 5474914
; GENERAL INFORMATION:
; APPLICANT: SPAETE, RICHARD
; TITLE OF INVENTION: METHOD OF INCREASING EXPRESSION
```

```
; TITLE OF INVENTION: OF VIRAL PROTEINS
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CHIRON CORPORATION
; STREET: 4560 Horton Street - R440
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608-2916
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/921,807B
; FILING DATE: 29-SEP-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MCCLUNG, BARBARA G.
; REGISTRATION NUMBER: 33,113
; REFERENCE/DOCKET NUMBER: 0209.001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 601-2708
; TELEFAX: (510) 655-3542
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 733 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-921-807B-6

Query Match          79.3%; Score 23; DB 1; Length 733;
Best Local Similarity 57.1%; Pred. No. 1.2e+03;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      1 VMXVAEF 7
Db      547 VMKIADF 553

RESULT 80
US-08-441-944A-6
; Sequence 6, Application US/08441944A
; Patent No. 5767250
; GENERAL INFORMATION:
; APPLICANT: SPAETE, RICHARD
; TITLE OF INVENTION: METHOD OF INCREASING EXPRESSION
; TITLE OF INVENTION: OF VIRAL PROTEINS
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CHIRON CORPORATION
; STREET: 4560 Horton Street - R440
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608-2916
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/441,944A
; FILING DATE: 16-MAY-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/921,807
; FILING DATE: 29-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: MCCLUNG, BARBARA G.
```

; REGISTRATION NUMBER: 33,113
; REFERENCE/DOCKET NUMBER: 0209.001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 601-2708
; TELEFAX: (510) 655-3542
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 733 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-441-944A-6

Query Match 79.3%; Score 23; DB 1; Length 733;
Best Local Similarity 57.1%; Pred. No. 1.2e+03;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VMXVAEF 7
|| :||
Db 547 VKIADF 553

RESULT 81
US-08-439-992A-4
; Sequence 4, Application US/08439992A
; Patent No. 6255454
; GENERAL INFORMATION:
; APPLICANT: Kiefer, Michael C.
; APPLICANT: Pablo, Valenzuela D.T.
; APPLICANT: Philip, Barr J.
; TITLE OF INVENTION: Expression and Use of Human Fibroblast
; TITLE OF INVENTION: Receptor
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30.
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/439,992A
; FILING DATE: 12-MAY-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Chung, Ling-Pong
; REGISTRATION NUMBER: 36,482
; REFERENCE/DOCKET NUMBER: 0165.004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 510-923-2704
; TELEFAX: (510) 655-3542
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 733 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-439-992A-4

Query Match 79.3%; Score 23; DB 3; Length 733;
Best Local Similarity 57.1%; Pred. No. 1.2e+03;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VMXVAEF 7
|| :||
Db 547 VKIADF 553

RESULT 82
US-08-471-570-8
; Sequence 8, Application US/08471570
; Patent No. 5750371
; GENERAL INFORMATION:
; APPLICANT: IGARASHI, Koichi
; APPLICANT: SENO, Masaharu
; APPLICANT: WATANABE, Tatsuya
; TITLE OF INVENTION: PROTEIN, DNA AND USE THEREOF
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS &
; STREET: 130 Water Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: US
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/471,570
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/149,664
; FILING DATE:
; APPLICATION NUMBER: US 07/743369
; FILING DATE: 16-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: LINEK, Ernest V
; REGISTRATION NUMBER: 29822
; REFERENCE/DOCKET NUMBER: 40897
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)523-3400
; TELEFAX: (617)523-6440
; TELEX: 200291 STRE UR
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 769 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-471-570-8

Query Match 79.3%; Score 23; DB 1; Length 769;
Best Local Similarity 57.1%; Pred. No. 1.2e+03;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VMXVAEF 7
|| :||
Db 640 VKIADF 646

RESULT 83
US-09-949-016-7119
; Sequence 7119, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7119
; LENGTH: 795
; TYPE: PRT
; ORGANISM: Human
; US-09-949-016-7119

Query Match 79.3%; Score 23; DB 4; Length 795;
Best Local Similarity 57.1%; Pred. No. 1.3e+03;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VMXVAEF 7
||:|:|

Db 619 VMKIADF 625

RESULT 84
US-09-383-630-6
; Sequence 6, Application US/09383630A
; Patent No. 6265632
; GENERAL INFORMATION:
; APPLICANT: Avner Yayon et al.
; TITLE OF INVENTION: ANIMAL MODEL FOR FIBROBLAST GROWTH
; FACTOR RECEPTOR ASSOCIATED
; CHONDRODYSPLASIA
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Mark M. Friedman c/o Anthony Castorina
; STREET: 2001 Jefferson Davis Highway, Suite 207
; CITY: Arlington
; STATE: Virginia
; COUNTRY: United States of America
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 1.44 megabyte, 3.5" microdisk
; COMPUTER: Twinhead* Slimnote-890TX
; OPERATING SYSTEM: MS DOS version 6.2,
; Windows version 3.11
; SOFTWARE: Word for Windows version 2.0 converted
; to an ASCII file
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/383,630A
; FILING DATE: 26-Aug-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: <Unknown>
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Friedman, Mark M.
; REGISTRATION NUMBER: 33,883
; REFERENCE/DOCKET NUMBER: 1402/2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 972-3-5625553
; TELEFAX: 972-3-5625554
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 801
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-383-630-6

Query Match 79.3%; Score 23; DB 3; Length 801;
Best Local Similarity 57.1%; Pred. No. 1.3e+03;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VMXVAEF 7
||:|:|

; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7119
; LENGTH: 795
; TYPE: PRT
; ORGANISM: Human
; US-09-949-016-7119

Query Match 79.3%; Score 23; DB 4; Length 795;
Best Local Similarity 57.1%; Pred. No. 1.3e+03;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VMXVAEF 7
||:|:|

Db 619 VMKIADF 625

RESULT 85
US-09-173-151A-33
; Sequence 33, Application US/09173151A
; Patent No. 6326472
; GENERAL INFORMATION:
; APPLICANT: Timans, Jacqueline C.
; APPLICANT: Debets, Johannes Eduard Maria
; APPLICANT: Antonius
; APPLICANT: Sana, Theodore R.
; APPLICANT: Bazan, J. Fernando
; APPLICANT: Kastelein, Robert A.
; TITLE OF INVENTION: Human Receptor Proteins; Related Reagents and Methods
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DNAX Research Institute
; STREET: 901 California Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/173,151A
; FILING DATE: 14-OCT-1998
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/065,776
; FILING DATE: 17-NOV-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/078,008
; FILING DATE: 12-MAR-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/081,883
; FILING DATE: 15-APR-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/095,987
; FILING DATE: 10-AUG-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/078,416
; FILING DATE: 18-MAR-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/062,066
; FILING DATE: 15-OCT-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ching, Edwin P.
; REGISTRATION NUMBER: 34,090
; REFERENCE/DOCKET NUMBER: DX0767X
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650)852-9196
; TELEFAX: (650)496-1200
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 802 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-173-151A-33

Query Match 79.3%; Score 23; DB 3; Length 802;
Best Local Similarity 57.1%; Pred. No. 1.3e+03;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VMXVAEF 7
||:|:|

Db 625 VKIADF 631

RESULT 86

US-09-383-630-3

Sequence 3, Application US/09383630A

Patent No. 6265632

GENERAL INFORMATION:

APPLICANT: Avner Yavon et al.

TITLE OF INVENTION: ANIMAL MODEL FOR FIBROBLAST GROWTH

FACTOR RECEPTOR ASSOCIATED

CHONDRODYSPLASIA

NUMBER OF SEQUENCES: 18

CORRESPONDENCE ADDRESS:

ADDRESSEE: Mark M. Friedman c/o Anthony Castorina

STREET: 2001 Jefferson Davis Highway, Suite 207

CITY: Arlington

STATE: Virginia

COUNTRY: United States of America

ZIP: 22202

COMPUTER READABLE FORM:

MEDIUM TYPE: 1.44 megabyte, 3.5" microdisk

COMPUTER: Twinhead* Slimnote-890TX

OPERATING SYSTEM: MS DOS version 6.2,

Windows version 3.11

SOFTWARE: Word for Windows version 2.0 converted

to an ASCII file

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/383,630A

FILING DATE: 26-Aug-1999

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: <Unknown>

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Friedman, Mark M.

REGISTRATION NUMBER: 33,883

REFERENCE/DOCKET NUMBER: 1402/2

TELECOMMUNICATION INFORMATION:

TELEPHONE: 972-3-5625553

TELEFAX: 972-3-5625554

TELEX: <Unknown>

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 806

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 3:

US-09-383-630-3

Query Match 79.3%; Score 23; DB 3; Length 806;

Best Local Similarity 57.1%; Pred. No. 1.3e+03;

Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7

Db 630 VKIADF 636

RESULT 87

US-07-640-029-1

Sequence 1, Application US/07640029

Patent No. 5229501

GENERAL INFORMATION:

APPLICANT: Kiefer, Michael C.

APPLICANT: Valenzuela, Pablo D.T.

APPLICANT: Barr, Philip J.

TITLE OF INVENTION: Expression and Use of Human Fibroblast

FACTOR RECEPTOR ASSOCIATED

CHONDRODYSPLASIA

NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:

ADDRESSEE: Chiron Corporation

STREET: 4560 Horton Street

CITY: Emeryville

STATE: California

COUNTRY: USA

ZIP: 94608

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/640,029

FILING DATE: 19910111

CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:

NAME: McClung, Barbara G.

REGISTRATION NUMBER: 33,113

REFERENCE/DOCKET NUMBER: CH-165

TELECOMMUNICATION INFORMATION:

TELEPHONE: 510-601-2708

TELEFAX: 510-655-3542

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 816 amino acids

TYPE: AMINO ACID

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-07-640-029-1

Query Match 79.3%; Score 23; DB 1; Length 816;

Best Local Similarity 57.1%; Pred. No. 1.3e+03;

Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7

Db 630 VKIADF 636

RESULT 88

US-09-949-016-10904

Sequence 10904, Application US/09949016

Patent No. 6812339

GENERAL INFORMATION:

APPLICANT: VENTER, J. Craig et al.

TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

FILE REFERENCE: CLO01307

CURRENT APPLICATION NUMBER: US/09/949,016

CURRENT FILING DATE: 2000-04-14

PRIOR APPLICATION NUMBER: 60/241,755

PRIOR FILING DATE: 2000-10-20

PRIOR APPLICATION NUMBER: 60/237,768

PRIOR FILING DATE: 2000-10-03

PRIOR APPLICATION NUMBER: 60/231,498

PRIOR FILING DATE: 2000-09-08

NUMBER OF SEQ ID NOS: 207012

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 10904

LENGTH: 816

TYPE: PRT

ORGANISM: Human

US-09-949-016-10904

Query Match 79.3%; Score 23; DB 4; Length 816;

Best Local Similarity 57.1%; Pred. No. 1.3e+03;

Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7

Db 639 VKIADF 645

RESULT 89
US-07-640-029-2
; Sequence 2, Application US/07640029
; Patent No. 5229501
; GENERAL INFORMATION:
; APPLICANT: Kiefer, Michael C.
; APPLICANT: Valenzuela, Pablo D.T.
; APPLICANT: Barr, Philip J.
; TITLE OF INVENTION: Expression and Use of Human Fibroblast
; TITLE OF INVENTION: Growth Factor Receptor
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: California
; COUNTRY: USA
; ZIP: 94608
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/640,029
; FILING DATE: 19910111
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: McClung, Barbara G.
; REGISTRATION NUMBER: 33,113
; REFERENCE/DOCKET NUMBER: CH-165
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 510-601-2708
; TELEFAX: 510-655-3542
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 817 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-07-640-029-2

Query Match 79.3%; Score 23; DB 1; Length 817;
Best Local Similarity 57.1%; Pred. No. 1.3e+03;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VMXVAEF 7
Db 632 VMKIADF 638

RESULT 90
US-07-921-807B-3
; Sequence 3, Application US/07921807B
; Patent No. 5474914
; GENERAL INFORMATION:
; APPLICANT: SPAETE, RICHARD
; TITLE OF INVENTION: METHOD OF INCREASING EXPRESSION
; TITLE OF INVENTION: OF VIRAL PROTEINS
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CHIRON CORPORATION
; STREET: 4560 Horton Street - R440
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608-2916
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/921,807B
; FILING DATE: 29-SEP-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: McCLUNG, BARBARA G.
; REGISTRATION NUMBER: 33,113
; REFERENCE/DOCKET NUMBER: 0209.001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 601-2708
; TELEFAX: (510) 655-3542
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 820 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-921-807B-3

Query Match 79.3%; Score 23; DB 1; Length 820;
Best Local Similarity 57.1%; Pred. No. 1.3e+03;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VMXVAEF 7
Db 634 VMKIADF 640

RESULT 91
US-08-441-944A-3
; Sequence 3, Application US/08441944A
; Patent No. 5767250
; GENERAL INFORMATION:
; APPLICANT: SPAETE, RICHARD
; TITLE OF INVENTION: METHOD OF INCREASING EXPRESSION
; TITLE OF INVENTION: OF VIRAL PROTEINS
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CHIRON CORPORATION
; STREET: 4560 Horton Street - R440
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608-2916
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/441,944A
; FILING DATE: 16-MAY-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/921,807
; FILING DATE: 29-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: McCLUNG, BARBARA G.
; REGISTRATION NUMBER: 33,113
; REFERENCE/DOCKET NUMBER: 0209.001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 601-2708
; TELEFAX: (510) 655-3542
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 820 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-441-944A-3

Query Match 79.3%; Score 23; DB 1; Length 820;
 Best Local Similarity 57.1%; Pred. No. 1.3e+03;
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
 || :||
 Db 634 VMKIADF 640

RESULT 92
 US-08-166-717D-6
 ; Sequence 6, Application US/08166717D
 ; Patent No. 5789182
 ; GENERAL INFORMATION:
 ; APPLICANT: Yavon, Avner
 ; APPLICANT: Ornitz, David M.
 ; APPLICANT: Klagsbrun, Michael
 ; APPLICANT: Leder, Philip
 ; TITLE OF INVENTION: SYSTEM FOR ASSAYING BINDING
 ; TITLE OF INVENTION: TO A HEPARIN-BINDING GROWTH
 ; TITLE OF INVENTION: FACTOR RECEPTOR
 ; NUMBER OF SEQUENCES: 6
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Clark & Elbing LLP
 ; STREET: 176 Federal Street
 ; CITY: Boston
 ; STATE: Massachusetts
 ; COUNTRY: U.S.A.
 ; ZIP: 02110

COMPUTER READABLE FORM:
 MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
 COMPUTER: IBM COMPATIBLE
 OPERATING SYSTEM: WINDOWS 95
 SOFTWARE: Wordperfect (Version 7.0)
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/166,717D
 FILING DATE: 12/14/93
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 07/631,717
 FILING DATE: 12/20/90
 ATTORNEY/AGENT INFORMATION:
 NAME: Kristina Bieker-Brady
 REGISTRATION NUMBER: 39,109
 REFERENCE/DOCKET NUMBER: 00383/017002
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617) 723-4123
 TELEFAX: (617) 723-8962
 TELEX:

INFORMATION FOR SEQ ID NO: 6:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 820
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 US-08-166-717D-6

Query Match 79.3%; Score 23; DB 1; Length 820;
 Best Local Similarity 57.1%; Pred. No. 1.3e+03;
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
 || :||
 Db 634 VMKIADF 640

RESULT 93
 US-08-439-992A-1
 ; Sequence 1, Application US/08439992A
 ; Patent No. 6255454
 ; GENERAL INFORMATION:
 ; APPLICANT: Kiefer, Michael C.
 ; APPLICANT: Pablo, Valenzuela D.T.

APPLICANT: Philip, Barr J.
 TITLE OF INVENTION: Expression and Use of Human Fibroblast
 TITLE OF INVENTION: Receptor
 NUMBER OF SEQUENCES: 12
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Chiron Corporation
 STREET: 4560 Horton Street
 CITY: Emeryville
 STATE: CA
 COUNTRY: USA
 ZIP: 94608

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/439,992A
 FILING DATE: 12-MAY-1995
 CLASSIFICATION: 530
 ATTORNEY/AGENT INFORMATION:
 NAME: Chung, Ling-Pong
 REGISTRATION NUMBER: 36,482
 REFERENCE/DOCKET NUMBER: 0165.004
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 510-923-2704
 TELEFAX: (510) 655-3542

INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 820 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-439-992A-1

Query Match 79.3%; Score 23; DB 3; Length 820;
 Best Local Similarity 57.1%; Pred. No. 1.3e+03;
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
 || :||
 Db 634 VMKIADF 640

RESULT 94
 US-08-451-822A-13
 ; Sequence 13, Application US/08451822A
 ; Patent No. 5863888
 ; GENERAL INFORMATION:
 ; APPLICANT: Dionne, Craig A
 ; APPLICANT: Crumley, Greg
 ; APPLICANT: Jays, Michael C
 ; APPLICANT: Schlessinger, Joseph
 ; TITLE OF INVENTION: Fibroblast Growth Factor Receptors
 ; NUMBER OF SEQUENCES: 19
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Rhone-Poulenc Rorer Legal Department
 ; STREET: 500 Arcola Road
 ; CITY: Collegeville
 ; STATE: PA
 ; COUNTRY: USA
 ; ZIP: 19426

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/451,822A
 FILING DATE: 26-MAY-1995
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/323,430
; FILING DATE: 14-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/934,372
; FILING DATE: 21-AUG-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/549,587
; FILING DATE: 06-JUL-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Savitzky, Martin
; REGISTRATION NUMBER: 29,699
; REFERENCE/DOCKET NUMBER: A0496E
; TELEPHONE: (610) 454-3816
; TELEFAX: (610) 454-3808
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 821 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-451-822A-13

Query Match 79.3%; Score 23; DB 2; Length 821;
Best Local Similarity 57.1%; Pred. No. 1.3e+03;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
DB 639 VMKIADF 645

RESULT 95
US-08-323-430-13
; Sequence 13, Application US/08323430
; Patent No. 6344546
; GENERAL INFORMATION:
; APPLICANT: Dionne, Craig A
; APPLICANT: Crumley, Greg
; APPLICANT: Jaye, Michael C
; APPLICANT: Schlessinger, Joseph
; TITLE OF INVENTION: Fibroblast Growth Factor Receptors
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Legal Department
; STREET: 500 Arcola Road
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/323,430
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US to be assigned
; FILING DATE: 21-AUG-1992
; APPLICATION NUMBER: US 07/549,587
; FILING DATE: 06-JUL-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Goodman, Rosanne
; REGISTRATION NUMBER: 32,534
; REFERENCE/DOCKET NUMBER: A0496
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 454-3817
; TELEFAX: (215) 454-3808
; INFORMATION FOR SEQ ID NO: 13:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 821 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-323-430-13

Query Match 79.3%; Score 23; DB 3; Length 821;
Best Local Similarity 57.1%; Pred. No. 1.3e+03;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
DB 639 VMKIADF 645

RESULT 96
US-07-997-133-1
; Sequence 1, Application US/07997133
; Patent No. 528855
; GENERAL INFORMATION:
; APPLICANT: Bergonzoni, Laura
; APPLICANT: Mazue, Guy
; APPLICANT: Isacchi, Antonella
; APPLICANT: Roncucci, Romeo
; TITLE OF INVENTION: Extracellular Form of the Human
; TITLE OF INVENTION: Fibroblast Growth Factor Receptor
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; STREET: 1755 Jefferson Davis Highway, Fourth Floor
; CITY: Arlington
; STATE: Virginia
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/997,133
; FILING DATE: 28-DEC-1992
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/642,755
; FILING DATE: 18-JAN-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Oblon, No. 528855man F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 769-226-0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 521-4500
; TELEFAX: (703) 486-2347
; TELEEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 822 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-997-133-1

Query Match 79.3%; Score 23; DB 1; Length 822;
Best Local Similarity 57.1%; Pred. No. 1.3e+03;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
DB 636 VMKIADF 642

```
RESULT 97
US-07-921-807B-4
; Sequence 4, Application US/07921807B
; Patent No. 5474914
; GENERAL INFORMATION:
; APPLICANT: SPAETE, RICHARD
; TITLE OF INVENTION: METHOD OF INCREASING EXPRESSION
; TITLE OF INVENTION: OF VIRAL PROTEINS
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CHIRON CORPORATION
; STREET: 4560 Horton Street - R440
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608-2916
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/07/921,807B
; FILING DATE: 29-SEP-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MCCLUNG, BARBARA G.
; REGISTRATION NUMBER: 33,113
; REFERENCE/DOCKET NUMBER: 0209.001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 601-2708
; TELEFAX: (510) 655-3542
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 822 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-921-807B-4
;
Query Match 79.3%; Score 23; DB 1; Length 822;
Best Local Similarity 57.1%; Pred. No. 1.3e+03;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
Db 636 VMKIADF 642

RESULT 99
US-08-441-944A-4
; Sequence 4, Application US/08441944A
; Patent No. 5767250
; GENERAL INFORMATION:
; APPLICANT: SPAETE, RICHARD
; TITLE OF INVENTION: METHOD OF INCREASING EXPRESSION
; TITLE OF INVENTION: OF VIRAL PROTEINS
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CHIRON CORPORATION
; STREET: 4560 Horton Street - R440
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608-2916
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 16-MAY-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/921,807
; FILING DATE: 29-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: MCCLUNG, BARBARA G.
; REGISTRATION NUMBER: 33,113
; REFERENCE/DOCKET NUMBER: 0209.001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 601-2708
; TELEFAX: (510) 655-3542
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 822 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-459-296-2
;
Query Match 79.3%; Score 23; DB 1; Length 822;
Best Local Similarity 57.1%; Pred. No. 1.3e+03;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
Db 636 VMKIADF 642

RESULT 98
US-08-459-296-2
; Sequence 2, Application US/08459296
; Patent No. 5670323
; GENERAL INFORMATION:
; APPLICANT: Gonzalez, Ana-Maria
; APPLICANT: Baird, Andrew
; TITLE OF INVENTION: PROCESS FOR DETECTION OF NEOPLASTIC
; TITLE OF INVENTION: DISEASE
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036/2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
```

US-08-441-944A-4

Job time : 33 secs

Query Match 79.3%; Score 23; DB 1; Length 822;
Best Local Similarity 57.1%; Pred. No. 1.3e+03;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VMXVAEF 7
Db 636 VMKIADF 642

RESULT 100

US-08-451-822A-12
; Sequence 12. Application US/08451822A
; Patent No. 5863888
; GENERAL INFORMATION:
; APPLICANT: Dionne, Craig A
; APPLICANT: Crumley, Greg
; APPLICANT: Jaye, Michael C
; APPLICANT: Schlessinger, Joseph
; TITLE OF INVENTION: Fibroblast Growth Factor Receptors
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Legal Department
; STREET: 500 Arcola Road
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/451,822A
; FILING DATE: 26-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/323,430
; FILING DATE: 14-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/934,372
; FILING DATE: 21-AUG-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/549,587
; FILING DATE: 06-JUL-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Savitzky, Martin
; REGISTRATION NUMBER: 29,699
; REFERENCE/DOCKET NUMBER: A0496E
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 454-3816
; TELEFAX: (610) 454-3808
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 822 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-451-822A-12

Query Match 79.3%; Score 23; DB 2; Length 822;
Best Local Similarity 57.1%; Pred. No. 1.3e+03;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VMXVAEF 7
Db 636 VMKIADF 642

Search completed: June 13, 2005, 14:00:49

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2005, 14:00:08 ; Search time 109 Seconds
(without alignments)
24.618 Million cell updates/sec

Title: 09730329-60ED

Perfect score: 29

Sequence: 1 VMXVAEF 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1710399 seqs, 383334425 residues

Total number of hits satisfying chosen parameters: 1710399

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 300 summaries

Database :

Published Applications AA:*

1:	/cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2:	/cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3:	/cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4:	/cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5:	/cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6:	/cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7:	/cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8:	/cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9:	/cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
10:	/cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
11:	/cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
12:	/cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13:	/cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
14:	/cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
15:	/cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
16:	/cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
17:	/cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
18:	/cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
19:	/cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep.*
20:	/cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
21:	/cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
22:	/cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	27	93.1	68	9	US-09-393-634-80
2	27	93.1	68	14	US-10-383-982-80
3	27	93.1	68	15	US-10-364-861-80
4	26	89.7	223	15	US-10-282-122A-63416
5	26	89.7	224	15	US-10-282-122A-64091
6	26	89.7	270	16	US-10-437-963-170879
7	25	86.2	406	15	US-10-282-122A-57253
8	24	82.8	118	16	US-10-425-115-187006
9	24	82.8	132	15	US-10-243-552-932
10	24	82.8	134	15	US-10-424-599-226832
11	24	82.8	145	13	US-10-013-379-11
					Sequence 80, Appl
					Sequence 80, Appl
					Sequence 63416, A
					Sequence 64091, A
					Sequence 170879, A
					Sequence 57253, A
					Sequence 187006, A
					Sequence 932, Appl
					Sequence 226832, A
					Sequence 11, Appl

12	82.8	180	15	US-10-424-599-170244	Sequence 170244, A
13	82.8	191	15	US-10-282-122A-52538	Sequence 52538, A
14	82.8	195	15	US-10-389-566-540	Sequence 540, Appl
15	82.8	195	15	US-10-282-122A-51903	Sequence 51903, A
16	82.8	238	15	US-10-282-122A-54190	Sequence 54190, A
17	82.8	335	9	US-09-815-242-13346	Sequence 13346, A
18	82.8	335	15	US-10-282-122A-73914	Sequence 73914, A
19	82.8	335	16	US-10-474-776-654	Sequence 654, Appl
20	82.8	335	17	US-10-472-928-1738	Sequence 1738, Appl
21	82.8	367	15	US-10-282-122A-51530	Sequence 51530, A
22	82.8	375	17	US-10-732-923-14226	Sequence 14226, A
23	82.8	378	16	US-10-739-930-5878	Sequence 5878, Appl
24	82.8	409	14	US-10-081-816-14	Sequence 14, Appl
25	82.8	419	16	US-10-437-963-131765	Sequence 131765, A
26	82.8	807	15	US-10-282-122A-46060	Sequence 46060, A
27	82.8	865	16	US-10-437-963-142919	Sequence 142919, A
28	82.8	902	15	US-10-282-122A-46068	Sequence 46068, A
29	82.8	919	15	US-10-369-493-5367	Sequence 5367, Appl
30	82.8	919	15	US-10-282-122A-70737	Sequence 70737, A
31	82.8	972	14	US-10-100-049-24	Sequence 24, Appl
32	82.8	1084	13	US-10-024-623-23	Sequence 23, Appl
33	82.8	1084	14	US-10-154-419-73	Sequence 73, Appl
34	82.8	1084	14	US-10-146-733-68	Sequence 68, Appl
35	82.8	1095	13	US-10-024-623-26	Sequence 26, Appl
36	82.8	1095	13	US-10-024-623-36	Sequence 36, Appl
37	82.8	1095	14	US-10-154-419-76	Sequence 76, Appl
38	82.8	1095	14	US-10-154-419-86	Sequence 86, Appl
39	82.8	1095	14	US-10-146-733-71	Sequence 71, Appl
40	82.8	1095	14	US-10-146-733-81	Sequence 81, Appl
41	82.8	1225	15	US-10-282-122A-56058	Sequence 56058, A
42	82.8	2059	16	US-10-437-963-200466	Sequence 200466, A
43	79.3	54	10	US-09-158-722-18	Sequence 18, Appl
44	79.3	54	10	US-09-158-722-46	Sequence 46, Appl
45	79.3	54	10	US-09-158-722-47	Sequence 47, Appl
46	79.3	61	15	US-10-424-599-169354	Sequence 169354, A
47	79.3	63	16	US-10-767-701-50646	Sequence 50646, A
48	79.3	68	11	US-09-864-408A-40	Sequence 40, Appl
49	79.3	90	9	US-09-393-634-76	Sequence 76, Appl
50	79.3	90	10	US-09-510-332-40	Sequence 40, Appl
51	79.3	90	14	US-10-383-982-76	Sequence 76, Appl
52	79.3	90	15	US-10-364-861-76	Sequence 76, Appl
53	79.3	90	16	US-10-770-127-40	Sequence 40, Appl
54	79.3	90	17	US-10-962-365-40	Sequence 40, Appl
55	79.3	122	15	US-10-424-599-173522	Sequence 173522, A
56	79.3	143	15	US-10-385-415-119	Sequence 119, Appl
57	79.3	150	9	US-09-815-242-4899	Sequence 4899, Appl
58	79.3	162	15	US-10-369-493-18481	Sequence 18481, A
59	79.3	173	16	US-10-437-963-184028	Sequence 184028, A
60	79.3	177	15	US-10-424-599-231564	Sequence 231564, A
61	79.3	189	14	US-10-238-075-1037	Sequence 1037, Appl
62	79.3	191	15	US-10-282-122A-77681	Sequence 77681, A
63	79.3	209	15	US-10-282-122A-72561	Sequence 72561, A
64	79.3	237	9	US-09-738-626-4458	Sequence 4458, Appl
65	79.3	237	16	US-10-767-701-36898	Sequence 36898, A
66	79.3	245	15	US-10-104-047-2492	Sequence 2492, Appl
67	79.3	245	15	US-10-334-143-182	Sequence 182, Appl
68	79.3	249	17	US-10-732-923-15746	Sequence 15746, A
69	79.3	251	17	US-10-732-923-15615	Sequence 15615, A
70	79.3	265	16	US-10-767-701-42888	Sequence 42888, A
71	79.3	274	16	US-10-470-957-23	Sequence 23, Appl
72	79.3	280	9	US-09-738-626-4505	Sequence 4505, Appl
73	79.3	287	15	US-10-424-599-147161	Sequence 147161, A
74	79.3	290	16	US-10-377-268-29	Sequence 29, Appl
75	79.3	293	15	US-10-282-122A-52955	Sequence 52955, A
76	79.3	299	16	US-10-334-143-203	Sequence 203, Appl
77	79.3	299	15	US-10-763-418-13	Sequence 13, Appl
78	79.3	300	16	US-10-763-418-31	Sequence 31, Appl
79	79.3	301	16	US-10-470-957-21	Sequence 21, Appl
80	79.3	309	15	US-10-369-493-11118	Sequence 11118, A
81	79.3	309	16	US-10-763-418-9	Sequence 9, Appl
82	79.3	310	9	US-09-939-754-7	Sequence 7, Appl
83	79.3	310	9	US-09-939-832-7	Sequence 7, Appl
84	79.3	310	9	US-09-939-833-7	Sequence 7, Appl

85	23	79.3	310	16	US-10-664-421-27	Sequence 27, Appl	158	789	17	US-10-659-004-2	Sequence 2, Appl
86	23	79.3	310	16	US-10-763-418-1	Sequence 1, Appl	159	789	17	US-10-659-004-4	Sequence 4, Appl
87	23	79.3	310	16	US-10-763-418-6	Sequence 6, Appl	160	789	17	US-10-659-004-6	Sequence 8, Appl
88	23	79.3	310	16	US-10-763-418-7	Sequence 7, Appl	161	789	17	US-10-659-004-8	Sequence 8, Appl
89	23	79.3	310	16	US-10-763-418-8	Sequence 8, Appl	162	789	17	US-10-659-004-10	Sequence 10, Appl
90	23	79.3	312	10	US-09-912-976-70	Sequence 70, Appl	163	789	17	US-10-659-004-12	Sequence 12, Appl
91	23	79.3	312	17	US-10-774-355A-2526	Sequence 2526, Ap	164	789	17	US-10-659-004-14	Sequence 14, Appl
92	23	79.3	314	10	US-09-510-332-44	Sequence 44, Appl	165	789	17	US-10-659-004-16	Sequence 16, Appl
93	23	79.3	314	10	US-09-510-332-71	Sequence 71, Appl	166	789	17	US-10-659-004-18	Sequence 18, Appl
94	23	79.3	314	14	US-10-017-161-1934	Sequence 1934, Ap	167	789	17	US-10-659-004-20	Sequence 20, Appl
95	23	79.3	314	15	US-10-292-798-1478	Sequence 1478, Ap	168	789	17	US-10-659-004-22	Sequence 22, Appl
96	23	79.3	314	15	US-10-343-650A-672	Sequence 672, App	169	789	17	US-10-659-004-24	Sequence 24, Appl
97	23	79.3	314	16	US-10-770-127-44	Sequence 44, Appl	170	789	17	US-10-659-004-26	Sequence 26, Appl
98	23	79.3	314	16	US-10-770-127-71	Sequence 71, Appl	171	789	17	US-10-659-004-28	Sequence 28, Appl
99	23	79.3	314	17	US-10-962-365-44	Sequence 44, Appl	172	789	17	US-10-659-004-30	Sequence 30, Appl
100	23	79.3	314	17	US-10-962-365-71	Sequence 71, Appl	173	789	17	US-10-659-004-32	Sequence 32, Appl
101	23	79.3	315	16	US-10-763-418-2	Sequence 2, Appl	174	789	17	US-10-659-004-34	Sequence 34, Appl
102	23	79.3	318	16	US-10-763-418-10	Sequence 10, Appl	175	789	17	US-10-659-004-36	Sequence 36, Appl
103	23	79.3	319	15	US-10-425-114-55531	Sequence 55531, A	176	789	17	US-10-659-004-38	Sequence 38, Appl
104	23	79.3	340	14	US-10-017-161-1822	Sequence 1822, Ap	177	802	9	US-09-758-386-3	Sequence 3, Appl
105	23	79.3	351	16	US-10-763-418-3	Sequence 3, Appl	178	802	14	US-10-011-548-33	Sequence 33, Appl
106	23	79.3	378	16	US-10-437-963-184025	Sequence 184025,	179	802	14	US-10-354-358-44	Sequence 44, Appl
107	23	79.3	379	16	US-10-664-421-45	Sequence 45, Appl	180	802	15	US-10-116-275-245	Sequence 245, App
108	23	79.3	383	15	US-10-282-122A-43909	Sequence 43909, A	181	802	15	US-10-394-322A-29	Sequence 29, Appl
109	23	79.3	383	17	US-10-857-625-727	Sequence 727, App	182	802	15	US-10-302-812-48	Sequence 48, Appl
110	23	79.3	387	9	US-09-815-242-5133	Sequence 5133, Ap	183	802	17	US-10-941-486-6	Sequence 6, Appl
111	23	79.3	387	15	US-10-282-122A-43554	Sequence 43554, A	184	806	15	US-10-394-322A-28	Sequence 28, Appl
112	23	79.3	393	15	US-10-369-493-21613	Sequence 21613, A	185	806	15	US-10-403-161-58	Sequence 58, Appl
113	23	79.3	394	13	US-10-087-192-1131	Sequence 1131, Ap	186	806	16	US-10-437-963-134077	Sequence 134077,
114	23	79.3	421	9	US-09-815-242-5303	Sequence 5303, Ap	187	806	17	US-10-696-639-42	Sequence 42, Appl
115	23	79.3	421	9	US-09-815-242-12388	Sequence 12388, A	188	806	17	US-10-941-486-7	Sequence 7, Appl
116	23	79.3	421	9	US-09-815-242-12820	Sequence 12820, A	189	808	15	US-10-403-161-60	Sequence 60, Appl
117	23	79.3	421	14	US-10-138-701-8	Sequence 8, Appl	190	811	16	US-10-437-963-152851	Sequence 152851,
118	23	79.3	421	14	US-10-138-701-22	Sequence 22, Appl	191	820	15	US-10-302-812-42	Sequence 42, Appl
119	23	79.3	421	15	US-10-282-122A-71223	Sequence 71223, A	192	820	17	US-10-482-029-164	Sequence 164, App
120	23	79.3	421	16	US-10-823-785-8	Sequence 8, Appl	193	821	15	US-10-394-322A-27	Sequence 27, Appl
121	23	79.3	421	16	US-10-823-785-22	Sequence 22, Appl	194	821	15	US-10-302-812-44	Sequence 44, Appl
122	23	79.3	423	15	US-10-282-122A-53283	Sequence 53283, A	195	821	17	US-10-941-486-8	Sequence 8, Appl
123	23	79.3	437	16	US-10-425-115-239771	Sequence 239771,	196	822	14	US-09-757-415A-2	Sequence 2, Appl
124	23	79.3	437	16	US-10-335-977-7239	Sequence 7239, Ap	197	822	14	US-10-204-041-2	Sequence 2, Appl
125	23	79.3	455	15	US-10-424-599-153508	Sequence 153508,	198	822	15	US-10-394-322A-26	Sequence 26, Appl
126	23	79.3	470	16	US-10-425-115-335184	Sequence 335184,	199	822	15	US-10-307-817-8	Sequence 8, Appl
127	23	79.3	477	15	US-10-425-114-43064	Sequence 43064, A	200	822	16	US-10-648-593-254	Sequence 254, App
128	23	79.3	480	15	US-10-425-114-63937	Sequence 63937, A	201	822	16	US-10-725-860-475	Sequence 475, App
129	23	79.3	483	17	US-10-732-923-2783	Sequence 2783, Ap	202	824	15	US-10-307-817-4	Sequence 4, Appl
130	23	79.3	484	15	US-10-369-493-6182	Sequence 6182, Ap	203	838	14	US-09-927-315-9	Sequence 9, Appl
131	23	79.3	484	17	US-10-732-923-2782	Sequence 2782, Ap	204	838	14	US-10-190-417-9	Sequence 9, Appl
132	23	79.3	509	16	US-10-425-115-205273	Sequence 205273,	205	838	17	US-10-679-102-9	Sequence 9, Appl
133	23	79.3	515	15	US-10-424-599-146129	Sequence 146129,	206	839	9	US-09-897-427A-4	Sequence 4, Appl
134	23	79.3	526	16	US-10-450-859-2	Sequence 2, Appl	207	839	14	US-10-035-045-21	Sequence 21, Appl
135	23	79.3	526	16	US-10-437-963-131085	Sequence 131085,	208	839	14	US-10-246-785-4	Sequence 4, Appl
136	23	79.3	530	15	US-10-425-114-60637	Sequence 60637, A	209	839	15	US-10-179-373-6	Sequence 6, Appl
137	23	79.3	531	15	US-10-425-114-45814	Sequence 45814, A	210	839	16	US-10-725-103-6	Sequence 6, Appl
138	23	79.3	560	16	US-10-425-115-335201	Sequence 335201,	211	839	16	US-10-725-489-6	Sequence 6, Appl
139	23	79.3	581	15	US-10-424-599-255268	Sequence 255268,	212	839	16	US-10-725-080A-6	Sequence 6, Appl
140	23	79.3	597	15	US-10-425-114-47187	Sequence 47187, A	213	839	16	US-10-725-472A-6	Sequence 6, Appl
141	23	79.3	612	17	US-10-732-923-13660	Sequence 13660, A	214	839	16	US-10-725-276-21	Sequence 21, Appl
142	23	79.3	650	14	US-10-193-477-119	Sequence 119, App	215	839	16	US-10-770-127-198	Sequence 198, App
143	23	79.3	650	15	US-10-403-847-116	Sequence 116, App	216	839	17	US-10-725-284-21	Sequence 21, Appl
144	23	79.3	682	17	US-10-868-577A-21	Sequence 21, Appl	217	839	17	US-10-725-418-6	Sequence 6, Appl
145	23	79.3	694	15	US-10-302-812-46	Sequence 46, Appl	218	839	17	US-10-679-102-9	Sequence 29, Appl
146	23	79.3	702	9	US-09-805-020-47	Sequence 47, Appl	219	911	15	US-10-282-122A-71564	Sequence 71564, A
147	23	79.3	735	15	US-10-307-817-6	Sequence 6, Appl	220	921	15	US-10-282-122A-60737	Sequence 60737, A
148	23	79.3	735	16	US-10-437-963-134076	Sequence 134076,	221	931	16	US-10-437-963-140717	Sequence 140717,
149	23	79.3	752	16	US-10-425-115-239772	Sequence 239772,	222	984	15	US-10-149-310-118	Sequence 118, App
150	23	79.3	759	17	US-10-941-486-9	Sequence 9, Appl	223	1002	17	US-10-732-923-13659	Sequence 13659, A
151	23	79.3	764	9	US-09-925-302-714	Sequence 714, App	224	1217	9	US-09-815-242-10627	Sequence 10627, A
152	23	79.3	764	10	US-09-925-302-714	Sequence 714, App	225	1217	15	US-10-282-122A-42405	Sequence 42405, A
153	23	79.3	773	9	US-09-815-242-11330	Sequence 11330, A	226	25	10	US-09-820-649-269	Sequence 269, App
154	23	79.3	773	15	US-10-282-122A-58722	Sequence 58722, A	227	25	14	US-10-160-162-269	Sequence 269, App
155	23	79.3	773	15	US-10-335-977-7240	Sequence 7240, Ap	228	25	17	US-10-936-773-269	Sequence 269, App
156	23	79.3	788	17	US-10-659-004-40	Sequence 40, Appl	229	37	15	US-10-425-114-49593	Sequence 49593, A
157	23	79.3	789	15	US-10-342-844-96	Sequence 96, Appl	230	61	10	US-09-978-303-19	Sequence 19, Appl

231	22	75.9	61	17	US-10-915-017-19	Sequence 19, Appl
232	22	75.9	66	16	US-10-425-115-212734	Sequence 212734, A
233	22	75.9	68	16	US-10-437-963-152970	Sequence 152970, A
234	22	75.9	72	14	US-10-437-963-152970	Sequence 29320, A
235	22	75.9	72	15	US-10-429-386-29320	Sequence 247856, A
236	22	75.9	90	16	US-10-429-386-29320	Sequence 119522, A
237	22	75.9	96	16	US-10-437-963-119522	Sequence 236103, A
238	22	75.9	102	16	US-10-425-115-236103	Sequence 190652, A
239	22	75.9	102	16	US-10-425-115-190652	Sequence 130652, A
240	22	75.9	112	14	US-10-238-075-1303	Sequence 1303, Ap
241	22	75.9	120	15	US-10-424-599-276785	Sequence 276785, A
242	22	75.9	134	9	US-09-805-020-65	Sequence 65, Appl
243	22	75.9	167	10	US-09-820-649-267	Sequence 267, App
244	22	75.9	167	14	US-10-160-162-267	Sequence 267, App
245	22	75.9	167	17	US-10-936-773-267	Sequence 267, App
246	22	75.9	168	16	US-10-767-701-34590	Sequence 34590, A
247	22	75.9	173	15	US-10-424-599-262081	Sequence 262081, A
248	22	75.9	177	16	US-10-425-115-366041	Sequence 366041, A
249	22	75.9	178	15	US-10-424-599-209172	Sequence 209172, A
250	22	75.9	179	14	US-10-166-698-8558	Sequence 8558, Ap
251	22	75.9	180	16	US-10-767-701-62544	Sequence 62544, A
252	22	75.9	184	15	US-10-424-599-177876	Sequence 177876, A
253	22	75.9	187	15	US-10-424-599-193330	Sequence 193330, A
254	22	75.9	191	9	US-09-860-670-110	Sequence 110, App
255	22	75.9	191	15	US-10-227-646-110	Sequence 110, App
256	22	75.9	197	16	US-10-451-467A-686	Sequence 686, App
257	22	75.9	230	15	US-10-264-049-2636	Sequence 2636, Ap
258	22	75.9	237	16	US-10-437-963-145305	Sequence 145305, A
259	22	75.9	255	16	US-10-425-115-298003	Sequence 298003, A
260	22	75.9	265	15	US-10-369-493-21568	Sequence 21568, A
261	22	75.9	275	15	US-10-282-122A-60512	Sequence 60512, A
262	22	75.9	282	16	US-10-377-268-32	Sequence 32, Appl
263	22	75.9	282	15	US-10-425-114-69980	Sequence 69980, A
264	22	75.9	289	16	US-10-767-701-40980	Sequence 40980, A
265	22	75.9	293	15	US-10-108-260A-2551	Sequence 2551, Ap
266	22	75.9	295	16	US-10-794-493-4	Sequence 4, Appl
267	22	75.9	295	16	US-10-794-493-5	Sequence 5, Appl
268	22	75.9	295	16	US-10-794-493-6	Sequence 6, Appl
269	22	75.9	295	16	US-10-794-493-7	Sequence 7, Appl
270	22	75.9	295	16	US-10-794-493-8	Sequence 8, Appl
271	22	75.9	310	15	US-10-282-122A-76370	Sequence 76370, A
272	22	75.9	312	16	US-10-664-421-24	Sequence 24, Appl
273	22	75.9	314	15	US-10-424-599-271724	Sequence 271724, A
274	22	75.9	331	16	US-10-425-115-328299	Sequence 328299, A
275	22	75.9	335	15	US-10-369-493-10345	Sequence 10345, A
276	22	75.9	340	16	US-10-664-421-43	Sequence 43, Appl
277	22	75.9	345	16	US-10-739-930-7323	Sequence 7323, Ap
278	22	75.9	347	9	US-09-738-626-6922	Sequence 6922, Ap
279	22	75.9	351	17	US-10-425-115-243901	Sequence 243901, A
280	22	75.9	353	17	US-10-732-923-13604	Sequence 13604, A
281	22	75.9	380	14	US-10-205-194-9	Sequence 9, Appl
282	22	75.9	380	17	US-10-739-930-7324	Sequence 7324, Ap
283	22	75.9	380	17	US-10-732-923-13448	Sequence 13448, A
284	22	75.9	381	15	US-10-282-122A-69946	Sequence 69946, A
285	22	75.9	388	15	US-10-282-122A-57497	Sequence 57497, A
286	22	75.9	395	15	US-10-369-493-20193	Sequence 20193, A
287	22	75.9	397	10	US-09-769-744A-152	Sequence 152, App
288	22	75.9	397	15	US-10-282-122A-74248	Sequence 74248, A
289	22	75.9	397	17	US-10-472-928-4698	Sequence 4698, Ap
290	22	75.9	402	15	US-10-282-122A-72509	Sequence 72509, A
291	22	75.9	418	15	US-10-282-122A-55471	Sequence 55471, A
292	22	75.9	430	16	US-10-437-963-127004	Sequence 127004, A
293	22	75.9	443	15	US-10-424-599-238107	Sequence 238107, A
294	22	75.9	458	16	US-10-425-115-311273	Sequence 311273, A
295	22	75.9	480	15	US-10-369-493-18824	Sequence 18824, A
296	22	75.9	500	14	US-10-128-714-8394	Sequence 8394, Ap
297	22	75.9	503	15	US-10-369-493-12878	Sequence 12878, A
298	22	75.9	519	16	US-10-425-115-239318	Sequence 239318, A
299	22	75.9	535	15	US-10-424-599-205550	Sequence 205550, A
300	22	75.9	543	13	US-10-108-605-257	Sequence 257, App

ALIGNMENTS

RESULT 1

US-09-393-634-80
; Sequence 80, Application US/09393634
; Patent No. US20020051997A1
; GENERAL INFORMATION:
; APPLICANT: Zuker, Charles S.
; APPLICANT: Adler, Jon Elliot
; APPLICANT: Ryba, Nick
; APPLICANT: Mueller, Ken
; APPLICANT: Hoon, Mark
; APPLICANT: The Regents of the University of California
; APPLICANT: The Government of the United States of America
; APPLICANT: as represented by the Secretary of the
; APPLICANT: Department of Health and Human Services
; TITLE OF INVENTION: SF, a No. US20020051997A1 Family of Taste Receptors
; FILE REFERENCE: 02307E-098000US
; CURRENT APPLICATION NUMBER: US/09/393,634
; CURRENT FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 92
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 80
; LENGTH: 68
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: human GR24
US-09-393-634-80

Query Match 93.1%; Score 27; DB 9; Length 68;
Best Local Similarity 71.4%; Pred. No. 30;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VMXVAEF 7
:| ||||
Db 11 IMVAEF 17

RESULT 2

US-10-383-982-80
; Sequence 80, Application US/10383982
; Publication No. US20030157568A1
; GENERAL INFORMATION:
; APPLICANT: Zuker, Charles S.
; APPLICANT: Ryba, Nick
; APPLICANT: Mueller, Ken
; APPLICANT: Hoon, Mark
; APPLICANT: The Regents of the University of California
; APPLICANT: The Government of the United States of America
; APPLICANT: as represented by the Secretary of the
; APPLICANT: Department of Health and Human Services
; TITLE OF INVENTION: SF, a No. US20030157568A1 Family of Taste Receptors
; FILE REFERENCE: 02307E-098000US
; CURRENT APPLICATION NUMBER: US/10/383,982
; CURRENT FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: US/09/393,634
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 92
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 80
; LENGTH: 68
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: human GR24
US-10-383-982-80

Query Match 93.1%; Score 27; DB 14; Length 68;
Best Local Similarity 71.4%; Pred. No. 30;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
QY      1 VMXVAEF 7
      :| |||||
Db      11 IMAVAEF 17

RESULT 3
US-10-364-861-80
; Sequence 80, Application US/10364861
; Publication No. US20040038312A1
; GENERAL INFORMATION:
; APPLICANT: Zuker, Charles S.
; APPLICANT: Adler, Jon Elliot
; APPLICANT: Hoon, Mark
; APPLICANT: Ryba, Nick
; APPLICANT: Mueller, Ken
; APPLICANT: The Regents of the University of California
; APPLICANT: The Government of the United States of America
; APPLICANT: as represented by the Secretary of the
; APPLICANT: Department of Health and Human Services
; TITLE OF INVENTION: T2R, a No. US20040038312A1e1 Family of Taste Receptors
; FILE REFERENCE: 02307E-098020US
; CURRENT APPLICATION NUMBER: US/10/364,861
; CURRENT FILING DATE: 2003-06-30
; PRIOR APPLICATION NUMBER: US 09/393,634
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 95
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 80
; LENGTH: 68
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: human T2R24, GR24 or SP24
US-10-364-861-80

Query Match      93.1%; Score 27; DB 15; Length 68;
Best Local Similarity 71.4%; Pred. No. 30;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 VMXVAEF 7
      :| |||||
Db      11 IMAVAEF 17

RESULT 4
US-10-282-122A-63416
; Sequence 63416, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zvekund, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614

QY      1 VMXVAEF 7
      :| |||||
Db      93 VMXVAEF 99

Query Match      89.7%; Score 26; DB 15; Length 223;
Best Local Similarity 71.4%; Pred. No. 1.8e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 VMXVAEF 7
      :| |||||
Db      93 VMXVAEF 99

RESULT 5
US-10-282-122A-64091
; Sequence 64091, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
```

```
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 64091
; LENGTH: 224
; TYPE: PRT
; ORGANISM: Mycoplasma pneumoniae
US-10-282-122A-64091

Query Match      89.7%; Score 26; DB 15; Length 224;
Best Local Similarity 71.4%; Pred. No. 1.8e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VMXVAEF 7
Db 93 VMNIAEF 99

RESULT 6
US-10-437-963-170879
; Sequence 170879, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 170879
; LENGTH: 770
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_69163C.1.pep
US-10-437-963-170879

Query Match      89.7%; Score 26; DB 16; Length 770;
Best Local Similarity 71.4%; Pred. No. 6.6e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VMXVAEF 7
Db 430 IMGVAEF 436

RESULT 7
US-10-282-122A-57253
; Sequence 57253, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: EUTRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
```

```
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 57253
; LENGTH: 406
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-10-282-122A-57253

Query Match      86.2%; Score 25; DB 15; Length 406;
Best Local Similarity 71.4%; Pred. No. 6.1e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VMXVAEF 7
Db 111 VLSVAEF 117

RESULT 8
US-10-425-115-187006
; Sequence 187006, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 187006
; LENGTH: 118
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(118)
; OTHER INFORMATION: unsure at all xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_102138C.1.pep
US-10-425-115-187006

Query Match      82.8%; Score 24; DB 16; Length 118;
Best Local Similarity 57.1%; Pred. No. 3e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VMXVAEF 7
Db 97 VLTIAEF 103
```

RESULT 9

US-10-243-552-932
; Sequence 932, Application US/10243552
; Publication No. US20030224379A1
; GENERAL INFORMATION:

; APPLICANT: Tang, Y. Tom
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Zhiwei
; APPLICANT: Wang, Gezhi
; APPLICANT: Ma, Yunging
; TITLE OF INVENTION: Novel Nucleic Acids and
; FILE REFERENCE: Polypeptides
; FILE REFERENCE: 807A

; CURRENT APPLICATION NUMBER: US/10/243,552
; CURRENT FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US 60/322,511
; PRIOR FILING DATE: 2001-09-13
; PRIOR APPLICATION NUMBER: PCT/US00/35017
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/488,725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: PCT/US01/02623
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: US 09/491,404
; PRIOR FILING DATE: 2000-01-25
; PRIOR APPLICATION NUMBER: PCT/US01/03800
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: US 09/496,914
; PRIOR FILING DATE: 2000-02-03
; PRIOR APPLICATION NUMBER: US 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: PCT/US01/04927
; PRIOR FILING DATE: 2001-02-26
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 998
; SOFTWARE: pt_FL_genes Version 5.0
; SEQ ID NO 932
; LENGTH: 132
; TYPE: PRT
; ORGANISM: Homo sapiens

US-10-243-552-932

Query Match 82.8%; Score 24; DB 15; Length 132;
Best Local Similarity 71.4%; Pred. No. 3.4e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VMXVAEF 7

Db 64 LNVVAEF 70

RESULT 10

US-10-424-599-226832
; Sequence 226832, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:

; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei

; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; Publication No. US20040031072A1
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 226832
; LENGTH: 134
; TYPE: PRT
; ORGANISM: Glycine max

; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(134)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_4685C.1.pep
US-10-424-599-226832

Query Match 82.8%; Score 24; DB 15; Length 134;
Best Local Similarity 71.4%; Pred. No. 3.4e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VMXVAEF 7

Db 64 VMGVAEF 70

RESULT 11

US-10-013-379-11
; Sequence 11, Application US/10013379
; Publication No. US20020188108A1
; GENERAL INFORMATION:

; APPLICANT: The Regents of the University of California
; APPLICANT: Cate, Jamie H.
; APPLICANT: No. US20020188108A1ler, Harry F.
; APPLICANT: Yusupov, Marat M.
; APPLICANT: Yusupova, Guinara ZH
; APPLICANT: Baucom, Albion
; APPLICANT: Lancaster, Laura
; APPLICANT: Dallas, Anne
; TITLE OF INVENTION: X-RAY CRYSTAL STRUCTURES OF FUNCTIONAL
; COMPLEXES OF THE BACTERIAL RIBOSOME CONTAINING TRANSFER RNA
; TITLE OF INVENTION: AND MODEL MESSENGER RNAs
; FILE REFERENCE: 196299-7010
; CURRENT APPLICATION NUMBER: US/10/013,379
; CURRENT FILING DATE: 2002-03-18
; PRIOR APPLICATION NUMBER: US 60/254,603
; PRIOR FILING DATE: 2000-12-09
; PRIOR APPLICATION NUMBER: US 60/278,013
; PRIOR FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: US 60/294,394
; PRIOR FILING DATE: 2001-05-30
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 145
; TYPE: PRT
; ORGANISM: Haloarcula marismortui

; FEATURE:
; OTHER INFORMATION: 50S ribosomal protein L13
; OTHER INFORMATION: lglyM
US-10-013-379-11

Query Match 82.8%; Score 24; DB 13; Length 145;
Best Local Similarity 83.3%; Pred. No. 3.7e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 MXVAEF 7

Db 1 MSVAEF 6

RESULT 12

US-10-424-599-170244
; Sequence 170244, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:

; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei

; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; Publication No. US20040031072A1
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 226832
; LENGTH: 134
; TYPE: PRT
; ORGANISM: Glycine max

FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/424,599
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 170244
LENGTH: 180
TYPE: PRT
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT3847_124745C.1.pgp
US-10-424-599-170244

Query Match 82.8%; Score 24; DB 15; Length 180;
Best Local Similarity 83.3%; Pred. No. 4.7e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 VMXVAEF 7
DB 116 MSVAEF 121

RESULT 13

US-10-282-122A-52538
Sequence 52538, Application US/10282122A
Publication No. US20040029129A1
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 52538
LENGTH: 191
TYPE: PRT
ORGANISM: Clostridium botulinum
US-10-282-122A-52538

Query Match 82.8%; Score 24; DB 15; Length 191;
Best Local Similarity 57.1%; Pred. No. 5e+02;

Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 1 VMXVAEF 7
DB 131 LMAIAEF 137

RESULT 14

US-10-389-566-540
Sequence 540, Application US/10389566
Publication No. US20040025202A1
GENERAL INFORMATION:
APPLICANT: Monsanto Technology, LLC
APPLICANT: Laurie, Cathy C
TITLE OF INVENTION: Nucleic Acid Molecules Associated with Oil in Plants
FILE REFERENCE: 38-77(52900)D
CURRENT APPLICATION NUMBER: US/10/389,566
CURRENT FILING DATE: 2003-03-31
PRIOR APPLICATION NUMBER: US 60/365,301
PRIOR FILING DATE: 2002-03-15
PRIOR APPLICATION NUMBER: US 60/391,786
PRIOR FILING DATE: 2002-06-25
PRIOR APPLICATION NUMBER: US 60/392,018
PRIOR FILING DATE: 2002-06-26
NUMBER OF SEQ ID NOS: 2459
SOFTWARE: PatentIn version 3.2
SEQ ID NO 540
LENGTH: 195
TYPE: PRT
ORGANISM: Glycine max
FEATURE:
NAME/KEY: misc feature
LOCATION: (132)-(132)
OTHER INFORMATION: Xaa can be any naturally occurring amino acid
US-10-389-566-540

Query Match 82.8%; Score 24; DB 15; Length 195;
Best Local Similarity 71.4%; Pred. No. 5.1e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
DB 73 VMLVAEY 79

RESULT 15

US-10-282-122A-51903
Sequence 51903, Application US/10282122A
Publication No. US20040029129A1
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335

```
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 51903
; LENGTH: 195
; TYPE: PRT
; ORGANISM: Clostridium acetobutylicum
US-10-282-122A-51903

Query Match      82.8%; Score 24; DB 15; Length 195;
Best Local Similarity 57.1%; Pred. No. 5.1e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      1 VMXVAEF 7
Db      131 LMAIAEF 137

RESULT 16
US-10-282-122A-54190
; Sequence 54190, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
```

```
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 54190
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Cornvebacterium diptheriae
US-10-282-122A-54190

Query Match      82.8%; Score 24; DB 15; Length 238;
Best Local Similarity 57.1%; Pred. No. 6.2e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      1 VMXVAEF 7
Db      72 VLSIAEF 78

RESULT 17
US-09-815-242-13346
; Sequence 13346, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13346
; LENGTH: 335
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-815-242-13346

Query Match      82.8%; Score 24; DB 9; Length 335;
Best Local Similarity 71.4%; Pred. No. 8.9e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 VMXVAEF 7
Db      224 VMSAAEF 230

RESULT 18
US-10-282-122A-73914
; Sequence 73914, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
```

```

; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 73914
; LENGTH: 335
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-282-122A-73914

Query Match      82.8%; Score 24; DB 15; Length 335;
Best Local Similarity 71.4%; Pred. No. 8.9e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 VMXVAEF 7
Db      224 VMSAAEF 230

RESULT 19
US-10-474-776-654
; Sequence 654, Application US/10474776
; Publication No. US20040110181A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; TITLE OF INVENTION: NOVEL STREPTOCOCCUS PNEUMONIAE OPEN READING FRAMES ENCODING POLYPEPTIDES
; FILE REFERENCE: AM100649-PCT
; CURRENT APPLICATION NUMBER: US/10/474,776
; CURRENT FILING DATE: 2003-10-14
; NUMBER OF SEQ ID NOS: 752
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 654
; LENGTH: 335
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-474-776-654

Query Match      82.8%; Score 24; DB 16; Length 335;
Best Local Similarity 71.4%; Pred. No. 8.9e+02;

APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 73914
; LENGTH: 335
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-282-122A-73914

Query Match      82.8%; Score 24; DB 17; Length 335;
Best Local Similarity 71.4%; Pred. No. 8.9e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 VMXVAEF 7
Db      224 VMSAAEF 230

RESULT 21
US-10-282-122A-51530
; Sequence 51530, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578

```

```

; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 51530
; LENGTH: 367
; TYPE: PRT
; ORGANISM: Clostridium acetobutylicum
US-10-282-122A-51530

Query Match      82.8%; Score 24; DB 15; Length 367;
Best Local Similarity 71.4%; Pred. No. 9.7e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 VMXVAEF 7
Db      29 VMEVAEY 35

RESULT 22
US-10-732-923-14226
; Sequence 14226, Application US/10732923
; Publication No. US20050108791A1
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 14226
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Anopheles gambiae str. PEST
US-10-732-923-14226

Query Match      82.8%; Score 24; DB 17; Length 375;
Best Local Similarity 83.3%; Pred. No. 1e+03;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 MXVAEF 7
Db      291 MTVAEF 296

RESULT 23
US-10-739-930-5878
; Sequence 5878, Application US/10739930
; Publication No. US20040216190A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
; FILE REFERENCE: 38-21(53377)B
; CURRENT APPLICATION NUMBER: US/10/739,930
; CURRENT FILING DATE: 2003-12-18
; NUMBER OF SEQ ID NOS: 11088
; SEQ ID NO 5878
; LENGTH: 378
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: Clone ID: ARATH-23APR03-C17092_1.p

; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 51530
; LENGTH: 367
; TYPE: PRT
; ORGANISM: Clostridium acetobutylicum
US-10-282-122A-51530

Query Match      82.8%; Score 24; DB 15; Length 367;
Best Local Similarity 71.4%; Pred. No. 9.7e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 VMXVAEF 7
Db      29 VMEVAEY 35

RESULT 22
US-10-732-923-14226
; Sequence 14226, Application US/10732923
; Publication No. US20050108791A1
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 14226
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Anopheles gambiae str. PEST
US-10-732-923-14226

Query Match      82.8%; Score 24; DB 17; Length 375;
Best Local Similarity 83.3%; Pred. No. 1e+03;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 MXVAEF 7
Db      291 MTVAEF 296

RESULT 23
US-10-739-930-5878
; Sequence 5878, Application US/10739930
; Publication No. US20040216190A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
; FILE REFERENCE: 38-21(53377)B
; CURRENT APPLICATION NUMBER: US/10/739,930
; CURRENT FILING DATE: 2003-12-18
; NUMBER OF SEQ ID NOS: 11088
; SEQ ID NO 5878
; LENGTH: 378
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: Clone ID: ARATH-23APR03-C17092_1.p
```

US-10-739-930-5878

Query Match 82.8%; Score 24; DB 16; Length 378;
Best Local Similarity 71.4%; Pred. No. 1e+03;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
Db 48 VMNMAEF 54

RESULT 24

US-10-081-816-14
; Sequence 14, Application US/10081816
; Publication No. US20030045472A1
; GENERAL INFORMATION:
; APPLICANT: Axel, Richard
; APPLICANT: Scott, Kristin
; TITLE OF INVENTION: Chemosensory Gene Family Encoding Gustatory And Olfactory Receptor
; FILE REFERENCE: 0575/64019-A/JPW/ADM
; CURRENT APPLICATION NUMBER: US/10/081,816
; CURRENT FILING DATE: 2002-02-22
; PRIOR APPLICATION NUMBER: 60/271,319
; PRIOR FILING DATE: 2001-02-23
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14
; LENGTH: 409
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-10-081-816-14

Query Match 82.8%; Score 24; DB 14; Length 409;
Best Local Similarity 71.4%; Pred. No. 1.1e+03;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
Db 24 VLLVAEF 30

RESULT 25

US-10-437-963-131765
; Sequence 131765, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 131765
; LENGTH: 419
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_337C.1.pep
US-10-437-963-131765

Query Match 82.8%; Score 24; DB 16; Length 419;
Best Local Similarity 71.4%; Pred. No. 1.1e+03;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VMXVAEF 7
|||
Db 94 VMLVAEY 100

RESULT 26

US-10-282-122A-46060
; Sequence 46060, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.

NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 46060
; LENGTH: 807
; TYPE: PRT
; ORGANISM: Bacillus anthracis

US-10-282-122A-46060

Query Match 82.8%; Score 24; DB 15; Length 807;
Best Local Similarity 83.3%; Pred. No. 2.2e+03;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 MXVAEF 7
|
Db 1 MTVAEF 6

RESULT 27

US-10-437-963-142919
; Sequence 142919, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei

; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 142919
; LENGTH: 865
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_43879C.1.pep
US-10-437-963-142919

Query Match 82.8%; Score 24; DB 16; Length 865;
Best Local Similarity 57.1%; Pred. No. 2.4e+03;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VMXVAEF 7
|||
Db 206 VMTIADF 212

RESULT 28

US-10-282-122A-46068
; Sequence 46068, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.

NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 46068
; LENGTH: 902

```
;
;
; TYPE: PRT
; ORGANISM: Bacillus anthracis
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (30)..(30)
; OTHER INFORMATION: X=any amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (215)..(215)
; OTHER INFORMATION: X=any amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (222)..(222)
; OTHER INFORMATION: X=any amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (543)..(543)
; OTHER INFORMATION: X=any amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (549)..(549)
; OTHER INFORMATION: X=any amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (556)..(556)
; OTHER INFORMATION: X=any amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (594)..(594)
; OTHER INFORMATION: X=any amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (597)..(597)
; OTHER INFORMATION: X=any amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (635)..(635)
; OTHER INFORMATION: X=any amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (649)..(649)
; OTHER INFORMATION: X=any amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (659)..(659)
; OTHER INFORMATION: X=any amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (663)..(663)
; OTHER INFORMATION: X=any amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (687)..(687)
; OTHER INFORMATION: X=any amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (693)..(693)
; OTHER INFORMATION: X=any amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (711)..(711)
; OTHER INFORMATION: X=any amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (718)..(718)
; OTHER INFORMATION: X=any amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (732)..(732)
; OTHER INFORMATION: X=any amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (737)..(737)
```

```
;
;
; OTHER INFORMATION: X=any amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (809)..(809)
; OTHER INFORMATION: X=any amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (810)..(810)
; OTHER INFORMATION: X=any amino acid
;
; US-10-282-122A-46069
```

Query Match 82.8%; Score 24; DB 15; Length 902;
Best Local Similarity 83.3%; Pred. No. 2.5e+03;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 MXVAEF 7
| | | | |
DB 113 MTVAEF 118

RESULT 29

```
US-10-369-493-5367
; Sequence 5367, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 5367
; LENGTH: 919
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
;
; US-10-369-493-5367
```

Query Match 82.8%; Score 24; DB 15; Length 919;
Best Local Similarity 71.4%; Pred. No. 2.5e+03;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
| | | | |
DB 592 VMIVAEY 598

RESULT 30

```
US-10-282-122A-70737
; Sequence 70737, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
```

```
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 70737
; LENGTH: 919
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-10-282-122A-70737

Query Match      82.8%; Score 24; DB 15; Length 919;
Best Local Similarity 83.3%; Pred. No. 2.5e+03;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 MXVAEF 7
DB      118 MSVAEF 123

RESULT 31
US-10-100-049-24
; Sequence 24, Application US/10100049
; Publication No. US20030078398A1
; GENERAL INFORMATION:
; APPLICANT: Graham, Margaret
; APPLICANT: Smith, Trevor
; APPLICANT: Munn, Edward
; APPLICANT: Knox, David
; APPLICANT: Oliver, Joanna
; APPLICANT: Newton, Susan
; TITLE OF INVENTION: RECOMBINANT DNA MOLECULES ENCODING AMINOPEPTIDASE ENZYMES
; TITLE OF INVENTION: AND THEIR USE IN THE PREPARATION OF VACCINES AGAINST HELMINTH
; TITLE OF INVENTION: INJECTIONS
; FILE REFERENCE: 1181-261
; CURRENT APPLICATION NUMBER: US/10/100,049
; CURRENT FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: US 09/129366
; PRIOR FILING DATE: 1998-08-05
; PRIOR APPLICATION NUMBER: US 08/335844
; PRIOR FILING DATE: 1995-01-09
; PRIOR APPLICATION NUMBER: PCT/GB93/00943
; PRIOR FILING DATE: 1993-05-07
; PRIOR APPLICATION NUMBER: GB 9209993.6
; PRIOR FILING DATE: 1992-05-08
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 24
; LENGTH: 972
; TYPE: PRT
; ORGANISM: Haemonchus contortus
US-10-100-049-24

Query Match      82.8%; Score 24; DB 14; Length 972;
Best Local Similarity 83.3%; Pred. No. 2.5e+03;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
Best Local Similarity 71.4%; Pred. No. 2.7e+03;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 VMXVAEF 7
DB      542 VISVAEF 548

RESULT 32
US-10-024-623-23
; Sequence 23, Application US/10024623
; Publication No. US20020187524A1
; GENERAL INFORMATION:
; APPLICANT: Curtis, Rory A.J.
; TITLE OF INVENTION: 8099, 46455, 54414, 53763, 67076, 67102, 44181,
; TITLE OF INVENTION: 67084FL, AND 67084 ALT, HUMAN PROTEINS AND METHODS OF
; TITLE OF INVENTION: USE THEREOF
; FILE REFERENCE: MNI-214CP
; CURRENT APPLICATION NUMBER: US/10/024,623
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US 60/256,240
; PRIOR FILING DATE: 2000-12-15
; PRIOR APPLICATION NUMBER: US 60/256,588
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: US 60/258,028
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 23
; LENGTH: 1084
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-024-623-23

Query Match      82.8%; Score 24; DB 13; Length 1084;
Best Local Similarity 71.4%; Pred. No. 3e+03;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 VMXVAEF 7
DB      1016 LMXVAEF 1022

RESULT 33
US-10-154-419-73
; Sequence 73, Application US/10154419
; Publication No. US20030143675A1
; GENERAL INFORMATION:
; APPLICANT: Curtis, Rory A.J.
; APPLICANT: Glucksmann, Maria Alexandra
; APPLICANT: Meyers, Rachel E.
; TITLE OF INVENTION: NOVEL 38594, 57312, 53659, 57250, 63760, 49938, 32146,
; TITLE OF INVENTION: 57259, 61118, 67067, 62092, 8099, 46455, 54414, 53763,
; TITLE OF INVENTION: 67076, 67102, 44181, 67084FL, 67084ALT, FBH58295FL, 57255,
; TITLE OF INVENTION: AND 57255alt MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-249
; CURRENT APPLICATION NUMBER: US/10/154,419
; CURRENT FILING DATE: 2002-05-22
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 99
; SOFTWARE: FastSEQ Version 4.0
; SEQ ID NO 73
; LENGTH: 1084
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-154-419-73

Query Match      82.8%; Score 24; DB 14; Length 1084;
Best Local Similarity 71.4%; Pred. No. 3e+03;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 VMXVAEF 7
DB      1016 LMXVAEF 1022
```

```
Db      1016 LMVVAEF 1022

RESULT 34
US-10-146-733-68
; Sequence 68, Application US/10146733
; Publication No. US20030165891A1
; GENERAL INFORMATION:
; APPLICANT: Curtis, Rory A.J.
; APPLICANT: Glucksmann, Maria A.
; TITLE OF INVENTION: NOVEL TWIK-6, TWIK-7, IC23927, TWIK-8, IC47611, IC47615,
; TITLE OF INVENTION: HNDA-1, TWIK-9, alpha2delta-4, 54414, AND 53763 MOLECULES
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: MNI-248
; CURRENT APPLICATION NUMBER: US/10/146,733
; CURRENT FILING DATE: 2002-05-15
; PRIOR APPLICATION NUMBER: US 60/185,938
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 09/518,520
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 09/518,866
; PRIOR FILING DATE: 2000-03-03
; PRIOR APPLICATION NUMBER: US 60/195,734
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: US 60/195,993
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/199,799
; PRIOR FILING DATE: 2000-04-26
; PRIOR APPLICATION NUMBER: US 60/233,537
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: US 60/235,018
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US 60/235,059
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US 60/256,240
; PRIOR FILING DATE: 2000-12-15
; PRIOR APPLICATION NUMBER: US 60/256,588
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: US 60/258,028
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: US 09/796,720
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 09/828,035
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: US 09/833,081
; PRIOR FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: US 09/843,128
; PRIOR FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: US 09/957,683
; PRIOR FILING DATE: 2001-09-19
; PRIOR APPLICATION NUMBER: US 09/964,252
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 09/964,256
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 10/024,623
; PRIOR FILING DATE: 2001-12-17
; NUMBER OF SEQ ID NOS: 85
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 68
; LENGTH: 1084
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-146-733-68

Query Match      82.8%; Score 24; DB 14; Length 1084;
Best Local Similarity 71.4%; Pred. No. 3e+03;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 VMXVAEF 7
Db      1016 LMVVAEF 1022

RESULT 35
US-10-024-623-26
; Sequence 26, Application US/10024623
; Publication No. US20020187524A1
; GENERAL INFORMATION:
; APPLICANT: Curtis, Rory A.J.
; TITLE OF INVENTION: 8099, 46455, 54414, 53763, 67076, 67102, 44181,
; TITLE OF INVENTION: 67084FL, AND 67084 ALT, HUMAN PROTEINS AND METHODS OF
; TITLE OF INVENTION: USE THEREOF
; FILE REFERENCE: MNI-214CP
; CURRENT APPLICATION NUMBER: US/10/024,623
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US 60/256,240
; PRIOR FILING DATE: 2000-12-15
; PRIOR APPLICATION NUMBER: US 60/256,588
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: US 60/258,028
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 26
; LENGTH: 1095
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-024-623-26

Query Match      82.8%; Score 24; DB 13; Length 1095;
Best Local Similarity 71.4%; Pred. No. 3e+03;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 VMXVAEF 7
Db      1016 LMVVAEF 1022

RESULT 36
US-10-024-623-36
; Sequence 36, Application US/10024623
; Publication No. US20020187524A1
; GENERAL INFORMATION:
; APPLICANT: Curtis, Rory A.J.
; TITLE OF INVENTION: 8099, 46455, 54414, 53763, 67076, 67102, 44181,
; TITLE OF INVENTION: 67084FL, AND 67084 ALT, HUMAN PROTEINS AND METHODS OF
; TITLE OF INVENTION: USE THEREOF
; FILE REFERENCE: MNI-214CP
; CURRENT APPLICATION NUMBER: US/10/024,623
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US 60/256,240
; PRIOR FILING DATE: 2000-12-15
; PRIOR APPLICATION NUMBER: US 60/256,588
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: US 60/258,028
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 36
; LENGTH: 1095
; TYPE: PRT
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 801, 1005
; OTHER INFORMATION: Xaa = any amino acid
US-10-024-623-36

Query Match      82.8%; Score 24; DB 13; Length 1095;
Best Local Similarity 71.4%; Pred. No. 3e+03;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 VMXVAEF 7
Db      1016 LMVVAEF 1022
```

```
RESULT 37
US-10-154-419-76
; Sequence 76, Application US/10154419
; Publication No. US20030143675A1
; GENERAL INFORMATION:
; APPLICANT: Curtis, Rory A.J.
; APPLICANT: Glucksmann, Maria A.
; TITLE OF INVENTION: NOVEL TWIK-6, TWIK-7, IC23927, TWIK-8, IC47611, IC47615,
; TITLE OF INVENTION: HMDA-1, TWIK-9, alpha2delta-4, 54414, AND 53763 MOLECULES
; TITLE OF INVENTION: AND USES THEREFOR
; FILE REFERENCE: MNI-249
; CURRENT APPLICATION NUMBER: US/10/154,419
; CURRENT FILING DATE: 2002-05-22
; PRIOR APPLICATION REMOVED - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 99
; SOFTWARE: FastSEQ Version 4.0
; SEQ ID NO 76
; LENGTH: 1095
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-154-419-76

Query Match      82.8%; Score 24; DB 14; Length 1095;
Best Local Similarity 71.4%; Pred. No. 3e+03;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 VMXVAEF 7
Db      1016 LMVVAEF 1022

RESULT 38
US-10-154-419-86
; Sequence 86, Application US/10154419
; Publication No. US20030143675A1
; GENERAL INFORMATION:
; APPLICANT: Curtis, Rory A.J.
; APPLICANT: Glucksmann, Maria Alexandra
; APPLICANT: Meyers, Rachel E.
; TITLE OF INVENTION: NOVEL 38594, 57312, 53659, 57250, 63760, 49938, 32146,
; TITLE OF INVENTION: 57259, 67118, 67067, 62092, 8099, 46455, 54414, 53763,
; TITLE OF INVENTION: 67076, 67102, 44181, 67084FL, 67084ALT, FBH58295FL, 57255,
; TITLE OF INVENTION: AND 57255salt MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-249
; CURRENT APPLICATION NUMBER: US/10/154,419
; CURRENT FILING DATE: 2002-05-22
; PRIOR APPLICATION REMOVED - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 99
; SOFTWARE: FastSEQ Version 4.0
; SEQ ID NO 86
; LENGTH: 1095
; TYPE: PRT
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 801, 1005
; OTHER INFORMATION: Xaa = any amino acid
US-10-154-419-86

Query Match      82.8%; Score 24; DB 14; Length 1095;
Best Local Similarity 71.4%; Pred. No. 3e+03;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 VMXVAEF 7
Db      1016 LMVVAEF 1022

RESULT 39
US-10-146-733-71
; Sequence 71, Application US/10146733
; Publication No. US20030165891A1
; GENERAL INFORMATION:
; APPLICANT: Curtis, Rory A.J.
; APPLICANT: Glucksmann, Maria A.
; TITLE OF INVENTION: NOVEL TWIK-6, TWIK-7, IC23927, TWIK-8, IC47611, IC47615,
; TITLE OF INVENTION: HMDA-1, TWIK-9, alpha2delta-4, 54414, AND 53763 MOLECULES
; TITLE OF INVENTION: AND USES THEREFOR
; FILE REFERENCE: MNI-249
; CURRENT APPLICATION NUMBER: US/10/146,733
; CURRENT FILING DATE: 2002-05-15
; PRIOR APPLICATION NUMBER: US 60/185,938
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 09/515,520
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 09/518,866
; PRIOR FILING DATE: 2000-03-03
; PRIOR APPLICATION NUMBER: US 60/195,734
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: US 60/195,993
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/199,799
; PRIOR FILING DATE: 2000-04-26
; PRIOR APPLICATION NUMBER: US 60/233,537
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: US 60/235,018
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US 60/235,059
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US 60/256,240
; PRIOR FILING DATE: 2000-12-15
; PRIOR APPLICATION NUMBER: US 60/256,588
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: US 60/258,028
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: US 09/796,720
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 09/828,035
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: US 09/833,081
; PRIOR FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: US 09/843,128
; PRIOR FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: US 09/957,683
; PRIOR FILING DATE: 2001-09-19
; PRIOR APPLICATION NUMBER: US 09/964,252
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 09/964,256
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 10/024,623
; PRIOR FILING DATE: 2001-12-17
; NUMBER OF SEQ ID NOS: 85
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 71
; LENGTH: 1095
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-146-733-71

Query Match      82.8%; Score 24; DB 14; Length 1095;
Best Local Similarity 71.4%; Pred. No. 3e+03;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 VMXVAEF 7
Db      1016 LMVVAEF 1022

RESULT 40
US-10-146-733-81
; Sequence 81, Application US/10146733
```

```
; Publication No. US20030165891A1
; GENERAL INFORMATION:
; APPLICANT: Curtis, Rory A.J.
; APPLICANT: Gluckmann, Maria A.
; APPLICANT: Silos-Santiago, Inmaculada
; TITLE OF INVENTION: NOVEL TWIK-6, TWIK-7, IC23927, TWIK-8, IC47611, IC47615,
; TITLE OF INVENTION: HNDA-1, TWIK-9, alpha2delta-4, 54414, AND 53763 MOLECULES
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: MNI-248
; CURRENT APPLICATION NUMBER: US/10/146,733
; CURRENT FILING DATE: 2002-05-15
; PRIOR APPLICATION NUMBER: US 60/185,938
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 09/515,520
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 09/518,866
; PRIOR FILING DATE: 2000-03-03
; PRIOR APPLICATION NUMBER: US 60/195,734
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: US 60/195,993
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/199,799
; PRIOR FILING DATE: 2000-04-26
; PRIOR APPLICATION NUMBER: US 60/233,537
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: US 60/235,018
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US 60/235,059
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US 60/256,240
; PRIOR FILING DATE: 2000-12-15
; PRIOR APPLICATION NUMBER: US 60/256,588
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: US 60/258,028
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: US 09/796,720
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 09/828,035
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: US 09/833,081
; PRIOR FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: US 09/843,128
; PRIOR FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: US 09/957,683
; PRIOR FILING DATE: 2001-09-19
; PRIOR APPLICATION NUMBER: US 09/964,252
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 09/964,256
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 10/024,623
; PRIOR FILING DATE: 2001-12-17
; NUMBER OF SEQ ID NOS: 85
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 81
; LENGTH: 1095
; TYPE: PRT
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 801..1005
; OTHER INFORMATION: Xaa = any amino acid
US-10-146-733-81
```

```
Query Match      82.8%; Score 24; DB 14; Length 1095;
Best Local Similarity 71.4%; Pred. No. 3e+03;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy      1 VMKVAEF 7
       :|:|:|:|
Db     1016 LMTVAEF 1022
```

RESULT 41

```
US-10-282-122A-56058
; Sequence 56058, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 56058
; LENGTH: 1225
; TYPE: PRT
; ORGANISM: Enterobacter cloacae
US-10-282-122A-56058
```

```
Query Match      82.8%; Score 24; DB 15; Length 1225;
Best Local Similarity 71.4%; Pred. No. 3.4e+03;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy      1 VMKVAEF 7
       |||:|:|
Db     945 VMEVAQF 951
```

RESULT 42

```
US-10-437-963-200466
; Sequence 200466, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
```

FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437,963
CURRENT FILING DATE: 2003-05-14
NUMBER OF SEQ ID NOS: 204966
SEQ ID NO 200466
LENGTH: 2059
TYPE: PRT
ORGANISM: Oryza sativa
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT4530_95932C.1.pep
US-10-437-963-200466

Query Match 82.8%; Score 24; DB 16; Length 2059;
Best Local Similarity 57.1%; Pred. No. 5.8e+03;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
DB 1869 ILAVAEF 1875

RESULT 43
US-09-158-722-18
Sequence 18, Application US/09158722
Publication No. US20030013848A1
GENERAL INFORMATION:
APPLICANT: Lemke Ph.D. et al., Greg E.
TITLE OF INVENTION: PROTEIN-TYROSINE KINASE GENES
NUMBER OF SEQUENCES: 54
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: CA
COUNTRY: US
ZIP: 92037

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
FILING DATE: 02-MAY-1994
APPLICATION NUMBER: US/09/158,722
CLASSIFICATION:
PRIOR APPLICATION DATA: US 08/456,647
FILING DATE: 02-JUN-1995
APPLICATION NUMBER: US 08/237,401
FILING DATE: 02-MAY-1994
PRIOR APPLICATION DATA:
FILING DATE: 15-MAY-1992
ATTORNEY/AGENT INFORMATION:
NAME: Wetherell Ph.D., John R.
REGISTRATION NUMBER: 31,678
REFERENCE/DOCKET NUMBER: 07251/007002
TELEPHONE: (619) 678-5070
TELEFAX: (619) 678-5099
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 54 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein

Query Match 79.3%; Score 23; DB 10; Length 54;
Best Local Similarity 57.1%; Pred. No. 2.4e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7

Db 9 VMKIADF 15

RESULT 44
US-09-158-722-46
Sequence 46, Application US/09158722
Publication No. US20030013848A1
GENERAL INFORMATION:
APPLICANT: Lemke Ph.D. et al., Greg E.
TITLE OF INVENTION: PROTEIN-TYROSINE KINASE GENES
NUMBER OF SEQUENCES: 54
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: CA
COUNTRY: US
ZIP: 92037

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
FILING DATE: 02-MAY-1994
APPLICATION NUMBER: US/09/158,722
CLASSIFICATION:
PRIOR APPLICATION DATA: US 08/456,647
FILING DATE: 02-JUN-1995
APPLICATION NUMBER: US 08/237,401
FILING DATE: 02-MAY-1994
PRIOR APPLICATION DATA:
FILING DATE: 15-MAY-1992
ATTORNEY/AGENT INFORMATION:
NAME: Wetherell Ph.D., John R.
REGISTRATION NUMBER: 31,678
REFERENCE/DOCKET NUMBER: 07251/007002
TELEPHONE: (619) 678-5070
TELEFAX: (619) 678-5099
INFORMATION FOR SEQ ID NO: 46:
SEQUENCE CHARACTERISTICS:
LENGTH: 54 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein

Query Match 79.3%; Score 23; DB 10; Length 54;
Best Local Similarity 57.1%; Pred. No. 2.4e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
DB 9 VMKIADF 15

RESULT 45
US-09-158-722-47
Sequence 47, Application US/09158722
Publication No. US20030013848A1
GENERAL INFORMATION:
APPLICANT: Lemke Ph.D. et al., Greg E.
TITLE OF INVENTION: PROTEIN-TYROSINE KINASE GENES
NUMBER OF SEQUENCES: 54
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla

```
; STATE: CA
; COUNTRY: US
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/158,722
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/456,647
; FILING DATE: 02-JUN-1995
; APPLICATION NUMBER: US 08/237,401
; FILING DATE: 02-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/884,486
; FILING DATE: 15-MAY-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Wetherell Ph.D., John R.
; REGISTRATION NUMBER: 31,678
; REFERENCE/DOCKET NUMBER: 07251/007002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 678-5070
; TELEFAX: (619) 678-5099
; INFORMATION FOR SEQ ID NO: 47:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 54 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-158-722-47

Query Match 79.3%; Score 23; DB 10; Length 54;
Best Local Similarity 57.1%; Pred. No. 2.4e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
DB 9 VMKIADF 15

RESULT 46
US-10-424-599-169354
; Sequence 169354, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 169354
; LENGTH: 61
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_123941C.1.pap
US-10-424-599-169354

Query Match 79.3%; Score 23; DB 15; Length 61;
Best Local Similarity 66.7%; Pred. No. 2.7e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 MXVAEF 7
```

```
DB 14 MTIAEF 19

RESULT 47
US-10-767-701-50646
; Sequence 50646, Application US/10767701
; Publication No. US20040172694A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 50646
; LENGTH: 63
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3480-021-Pl-K1-H3.pap
US-10-767-701-50646

Query Match 79.3%; Score 23; DB 16; Length 63;
Best Local Similarity 71.4%; Pred. No. 2.8e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
DB 24 VMQTAEF 30

RESULT 48
US-09-864-408A-40
; Sequence 40, Application US/09864408A
; Publication No. US20040009474A1
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; APPLICANT: Shimkets, Richard A.
; TITLE OF INVENTION: No. US20040009474A1 Human Polynucleotides and Polypeptides Enco
; FILE REFERENCE: 21402-012
; CURRENT APPLICATION NUMBER: US/09/864,408A
; CURRENT FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 60/206,690
; PRIOR FILING DATE: 2000-05-24
; NUMBER OF SEQ ID NOS: 9068
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 40
; LENGTH: 68
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-864-408A-40

Query Match 79.3%; Score 23; DB 11; Length 68;
Best Local Similarity 83.3%; Pred. No. 3e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 MXVAEF 7
DB 1 MLVAEF 6

RESULT 49
US-09-393-634-76
; Sequence 76, Application US/09393634
; Patent No. US20020051997A1
; GENERAL INFORMATION:
; APPLICANT: Zuker, Charles S.
; APPLICANT: Adler, Jon Elliot
; APPLICANT: Ryba, Nick
```


; APPLICANT: Mueller, Ken
; APPLICANT: Hoon, Mark
; APPLICANT: The Regents of the University of California
; APPLICANT: The Government of the United States of America
; APPLICANT: as represented by the Secretary of the
; APPLICANT: Department of Health and Human Services
; TITLE OF INVENTION: SF, a No. US20020051997A1el Family of Taste Receptors
; FILE REFERENCE: 02307E-098000US
; CURRENT APPLICATION NUMBER: US/09/393,634
; CURRENT FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 92
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 76
; LENGTH: 90
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: human GR22
; NAME/KEY: MOD_RES
; LOCATION: (1)..(90)
; OTHER INFORMATION: Xaa = any amino acid
US-09-393-634-76

Query Match 79.3%; Score 23; DB 9; Length 90;
Best Local Similarity 42.9%; Pred. No. 4.1e+02;
Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VMXVAEF 7
Db 11 ILAIAEF 17

RESULT 50
US-09-510-332-40
; Sequence 40, Application US/09510332
; Publication No. US20030022278A1
; GENERAL INFORMATION:
; APPLICANT: Zuker, Charles S.
; APPLICANT: Adler, Jon Elliot
; APPLICANT: Ryba, Nick
; APPLICANT: Mueller, Ken
; APPLICANT: Hoon, Mark
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: T2R, a No. US20030022278A1el Family of Taste Receptors
; FILE REFERENCE: 02307E-098010US
; CURRENT APPLICATION NUMBER: US/09/510,332
; CURRENT FILING DATE: 2000-02-22
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 172
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 40
; LENGTH: 90
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: human T2R22 (hGR22)
; NAME/KEY: MOD_RES
; LOCATION: (1)..(90)
; OTHER INFORMATION: Xaa = any amino acid
US-09-510-332-40

Query Match 79.3%; Score 23; DB 10; Length 90;
Best Local Similarity 42.9%; Pred. No. 4.1e+02;
Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VMXVAEF 7
Db 11 ILAIAEF 17

RESULT 51
US-10-383-982-76

; Sequence 76, Application US/10383982
; Publication No. US20030157568A1
; GENERAL INFORMATION:
; APPLICANT: Zuker, Charles S.
; APPLICANT: Adler, Jon Elliot
; APPLICANT: Ryba, Nick
; APPLICANT: Mueller, Ken
; APPLICANT: Hoon, Mark
; APPLICANT: The Regents of the University of California
; APPLICANT: The Government of the United States of America
; APPLICANT: as represented by the Secretary of the
; APPLICANT: Department of Health and Human Services
; TITLE OF INVENTION: SF, a No. US20030157568A1el Family of Taste Receptors
; FILE REFERENCE: 02307E-098000US
; CURRENT APPLICATION NUMBER: US/10/383,982
; CURRENT FILING DATE: 2003-03-07
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 92
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 76
; LENGTH: 90
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (1)..(90)
; OTHER INFORMATION: human GR22
; OTHER INFORMATION: Xaa = any amino acid
US-10-383-982-76

Query Match 79.3%; Score 23; DB 14; Length 90;
Best Local Similarity 42.9%; Pred. No. 4.1e+02;
Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VMXVAEF 7
Db 11 ILAIAEF 17

RESULT 52
US-10-364-861-76
; Sequence 76, Application US/10364861
; Publication No. US20040038312A1
; GENERAL INFORMATION:
; APPLICANT: Zuker, Charles S.
; APPLICANT: Adler, Jon Elliot
; APPLICANT: Hoon, Mark
; APPLICANT: Ryba, Nick
; APPLICANT: Mueller, Ken
; APPLICANT: The Regents of the University of California
; APPLICANT: The Government of the United States of America
; APPLICANT: as represented by the Secretary of the
; APPLICANT: Department of Health and Human Services
; TITLE OF INVENTION: T2R, a No. US20040038312A1el Family of Taste Receptors
; FILE REFERENCE: 02307E-098020US
; CURRENT APPLICATION NUMBER: US/10/364,861
; CURRENT FILING DATE: 2003-06-30
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 95
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 76
; LENGTH: 90
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: human T2R22, GR22 or SF22
; NAME/KEY: MOD_RES
; LOCATION: (1)..(90)
; OTHER INFORMATION: Xaa = any amino acid
US-10-364-861-76

Query Match 79.3%; Score 23; DB 15; Length 90;
Best Local Similarity 42.9%; Pred. No. 4.1e+02;
Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
:: :|||
Db 11 ILAIAEF 17

RESULT 53
US-10-770-127-40
; Sequence 40, Application US/10770127
; Publication No. US20040214239A1
; GENERAL INFORMATION:
; APPLICANT: SERVANT, GUY
; APPLICANT: OZECK, MARK
; APPLICANT: BRUST, PAUL
; APPLICANT: XU, HONG
; TITLE OF INVENTION: FUNCTIONAL COUPLING OF T1RS AND T2RS BY GI PROTEINS
; TITLE OF INVENTION: AND CELL-BASED ASSAYS FOR THE IDENTIFICATION OF T1R
; FILE REFERENCE: 100337.54281US
; CURRENT FILING DATE: 2004-02-03
; PRIOR FILING DATE: 2004-02-03
; PRIOR FILING DATE: 2003-02-03
; PRIOR APPLICATION NUMBER: 60/444,172
; PRIOR FILING DATE: 2003-02-03
; PRIOR APPLICATION NUMBER: 60/457,318
; PRIOR FILING DATE: 2003-03-26
; NUMBER OF SEQ ID NOS: 210
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 40
; LENGTH: 90
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MOD RES
; LOCATION: (35)..(35)
; OTHER INFORMATION: Variable amino acid
; FEATURE:
; NAME/KEY: MOD RES
; LOCATION: (39)..(39)
; OTHER INFORMATION: Variable amino acid
; FEATURE:
; NAME/KEY: MOD RES
; LOCATION: (82)..(82)
; OTHER INFORMATION: Variable amino acid
US-10-770-127-40

Query Match 79.3%; Score 23; DB 16; Length 90;
Best Local Similarity 42.9%; Pred. No. 4.1e+02;
Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
:: :|||
Db 11 ILAIAEF 17

RESULT 54
US-10-962-365-40
; Sequence 40, Application US/10962365
; Publication No. US20050048586A1
; GENERAL INFORMATION:
; APPLICANT: Zuker, Charles S.
; APPLICANT: Adler, Jon Elliot
; APPLICANT: Ryba, Nick
; APPLICANT: Mueller, Ken
; APPLICANT: Hoon, Mark
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: T2R, a Novel Family of Taste Receptors
; FILE REFERENCE: 02307E-098010US
; CURRENT FILING DATE: 2004-10-07
; PRIOR APPLICATION NUMBER: US/10/962,365
; CURRENT FILING DATE: 2004-10-07

; PRIOR APPLICATION NUMBER: US/09/510,332
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 09/393,634
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 172
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 40
; LENGTH: 90
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MOD RES
; LOCATION: (1)..(90)
; OTHER INFORMATION: human T2R22 (hGR22)
; OTHER INFORMATION: Xaa = any amino acid
US-10-962-365-40

Query Match 79.3%; Score 23; DB 17; Length 90;
Best Local Similarity 42.9%; Pred. No. 4.1e+02;
Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
:: :|||
Db 11 ILAIAEF 17

RESULT 55
US-10-424-599-173522
; Sequence 173522, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David X
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 173522
; LENGTH: 122
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT MRT3847_127708C.1.pap
US-10-424-599-173522

Query Match 79.3%; Score 23; DB 15; Length 122;
Best Local Similarity 83.3%; Pred. No. 5.6e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 MKVAEF 7
| :|||
Db 34 MKVAEF 39

RESULT 56
US-10-385-415-119
; Sequence 119, Application US/10385415
; Publication No. US20040014158A1
; GENERAL INFORMATION:
; APPLICANT: Bacher, Adelbert
; APPLICANT: Fischer, Markus
; TITLE OF INVENTION: PROTEIN CONJUGATES, METHOD, VECTORS, PROTEINS AND DNA FOR
; TITLE OF INVENTION: PRODUCING THEM, THEIR USE AND MEDICAMENTS AND VACCINES CONTAINING
; TITLE OF INVENTION: A CERTAIN QUANTITY OF SAID PROTEIN CONJUGATES
; FILE REFERENCE: 9286.6CT
; CURRENT APPLICATION NUMBER: US/10/385,415
; CURRENT FILING DATE: 2003-03-10
; PRIOR APPLICATION NUMBER: US 09/936,028
; PRIOR FILING DATE: 2002-01-31

;/ PRIOR APPLICATION NUMBER: PCT/EP00/01899
;/ PRIOR FILING DATE: 2000-03-03
;/ PRIOR APPLICATION NUMBER: DE 19910102.7
;/ PRIOR FILING DATE: 1999-03-08
;/ NUMBER OF SEQ ID NOS: 154
;/ SOFTWARE: PatentIn version 3.2
;/ SEQ ID NO 119
;/ LENGTH: 143
;/ TYPE: PRT
;/ ORGANISM: Archaeoglobus fulgidus
US-10-385-415-119

Query Match 79.3%; Score 23; DB 15; Length 143;
Best Local Similarity 83.3%; Pred. No. 6.6e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 MXVAEF 7
| | | | |
Db 8 MVVAEF 13

RESULT 57
US-09-815-242-4899
;/ Sequence 4899, Application US/09815242
;/ Patent No. US2002061569A1
;/ GENERAL INFORMATION:
;/ APPLICANT: Haselbeck, Robert
;/ APPLICANT: Ohlsen, Kari L.
;/ APPLICANT: Zyskind, Judith W.
;/ APPLICANT: Wall, Daniel
;/ APPLICANT: Trawick, John D.
;/ APPLICANT: Carr, Grant J.
;/ APPLICANT: Yamamoto, Robert T.
;/ APPLICANT: Xu, H. Howard
;/ TITLE OF INVENTION: Identification of Essential Genes in
;/ FILE OF INVENTION: Prokaryotes
;/ FILE REFERENCE: ELITRA.011A
;/ CURRENT APPLICATION NUMBER: US/09/815,242
;/ CURRENT FILING DATE: 2001-03-21
;/ PRIOR APPLICATION NUMBER: 60/191,078
;/ PRIOR FILING DATE: 2000-03-21
;/ PRIOR APPLICATION NUMBER: 60/206,848
;/ PRIOR FILING DATE: 2000-05-23
;/ PRIOR APPLICATION NUMBER: 60/207,727
;/ PRIOR FILING DATE: 2000-05-26
;/ PRIOR APPLICATION NUMBER: 60/242,578
;/ PRIOR FILING DATE: 2000-10-23
;/ PRIOR APPLICATION NUMBER: 60/253,625
;/ PRIOR FILING DATE: 2000-11-27
;/ PRIOR APPLICATION NUMBER: 60/257,931
;/ PRIOR FILING DATE: 2000-12-22
;/ PRIOR APPLICATION NUMBER: 60/269,308
;/ PRIOR FILING DATE: 2001-02-16
;/ NUMBER OF SEQ ID NOS: 14110
;/ SOFTWARE: FastSeq for Windows Version 4.0
;/ SEQ ID NO 4899
;/ LENGTH: 150
;/ TYPE: PRT
;/ ORGANISM: Enterococcus faecalis
US-09-815-242-4899

Query Match 79.3%; Score 23; DB 9; Length 150;
Best Local Similarity 57.1%; Pred. No. 6.9e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VMXVAEF 7
| | | | |
Db 40 LMDIAEF 46

RESULT 58
US-10-369-493-18481
;/ Sequence 18481, Application US/10369493

;/ Publication No. US20030233675A1
;/ GENERAL INFORMATION:
;/ APPLICANT: Cao, Yongwei
;/ APPLICANT: Hinkle, Gregory J.
;/ APPLICANT: Slater, Steven C.
;/ APPLICANT: Goldman, Barry S.
;/ APPLICANT: Chen, Xianfeng
;/ TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
;/ FILE REFERENCE: 38-10(52052)B
;/ CURRENT APPLICATION NUMBER: US/10/369,493
;/ CURRENT FILING DATE: 2003-02-28
;/ PRIOR APPLICATION NUMBER: US 60/360,039
;/ PRIOR FILING DATE: 2002-02-21
;/ NUMBER OF SEQ ID NOS: 47374
;/ SEQ ID NO 18481
;/ LENGTH: 162
;/ TYPE: PRT
;/ ORGANISM: Lactococcus lactis
US-10-369-493-18481

Query Match 79.3%; Score 23; DB 15; Length 162;
Best Local Similarity 66.7%; Pred. No. 7.5e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 MXVAEF 7
| | | | |
Db 37 MTIAEF 42

RESULT 59
US-10-437-963-184028
;/ Sequence 184028, Application US/10437963
;/ Publication No. US20040123343A1
;/ GENERAL INFORMATION:
;/ APPLICANT: La Rosa, Thomas J.
;/ APPLICANT: Kovalic, David K.
;/ APPLICANT: Zhou, Yihua
;/ APPLICANT: Cao, Yongwei
;/ APPLICANT: Wu, Wei
;/ APPLICANT: Boukharov, Andrey A.
;/ APPLICANT: Barbazuk, Brad
;/ APPLICANT: Li, Ping
;/ TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
;/ FILE REFERENCE: 38-21(53221)B
;/ CURRENT APPLICATION NUMBER: US/10/437,963
;/ CURRENT FILING DATE: 2003-05-14
;/ NUMBER OF SEQ ID NOS: 204966
;/ SEQ ID NO 184028
;/ LENGTH: 173
;/ TYPE: PRT
;/ ORGANISM: Oryza sativa
;/ FEATURE:
;/ OTHER INFORMATION: Clone ID: PAT_MRT4530_81060C.1.pep
US-10-437-963-184028

Query Match 79.3%; Score 23; DB 16; Length 173;
Best Local Similarity 83.3%; Pred. No. 8e+02; 1; Indels 0; Gaps 0;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 MXVAEF 7
| | | | |
Db 92 MGVAEF 97

RESULT 60
US-10-424-599-231564
;/ Sequence 231564, Application US/10424599
;/ Publication No. US20040031072A1
;/ GENERAL INFORMATION:
;/ APPLICANT: La Rosa Thomas J
;/ APPLICANT: Kovalic David K

; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 231564
; LENGTH: 177
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_51123C.1.pep
US-10-424-599-231564

Query Match 79.3%; Score 23; DB 15; Length 177;
Best Local Similarity 57.1%; Pred. No. 8.2e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
Db 84 MMLIAEF 90

RESULT 61
US-10-238-075-1037
; Sequence 1037, Application US/10238075
; Publication No. US20030148324A1
; GENERAL INFORMATION:
; APPLICANT: I.N.S.E.R.M.

; TITLE OF INVENTION: Polynucleotides which are of nature B2/D- A- and which are isolated
; FILE REFERENCE: BLANDINE
; CURRENT APPLICATION NUMBER: US/10/238,075
; CURRENT FILING DATE: 2002-09-10
; PRIOR APPLICATION NUMBER: 0003145
; PRIOR FILING DATE: 2000-03-10
; NUMBER OF SEQ ID NOS: 1576
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1037
; LENGTH: 189
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-238-075-1037

Query Match 79.3%; Score 23; DB 14; Length 189;
Best Local Similarity 57.1%; Pred. No. 8.7e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
Db 23 ILDVAEF 29

RESULT 62
US-10-282-122A-77681
; Sequence 77681, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu

; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 77681
; LENGTH: 191
; TYPE: PRT
; ORGANISM: Vibrio cholerae
US-10-282-122A-77681

Query Match 79.3%; Score 23; DB 15; Length 191;
Best Local Similarity 83.3%; Pred. No. 8.8e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 MXVAEF 7
Db 1 MRVAEF 6

RESULT 63
US-10-282-122A-72561
; Sequence 72561, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578

```
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 72561
; LENGTH: 209
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-10-282-122A-72561

Query Match          79.3%; Score 23; DB 15; Length 209;
Best Local Similarity 57.1%; Pred. No. 9.7e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy      1 VMXVAEF 7
Db      64 IISVAEF 70

RESULT 64
US-09-738-626-4458
; Sequence 4458, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OKAZI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 4458
; LENGTH: 237
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-4458

Query Match          79.3%; Score 23; DB 9; Length 237;
Best Local Similarity 42.9%; Pred. No. 1.1e+03;
Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy      1 VMXVAEF 7
Db      74 ILSIAEF 80

RESULT 65
US-10-767-701-36898
; Sequence 36898, Application US/10767701
; Publication No. US20040172684A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 36898
; LENGTH: 237
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)...(237)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C120919_1.pep
US-10-767-701-36898

Query Match          79.3%; Score 23; DB 16; Length 237;
Best Local Similarity 85.7%; Pred. No. 1.1e+03;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 VMXVAEF 7
Db      136 VMXVGEF 142

RESULT 66
US-10-104-047-2492
; Sequence 2492, Application US/10104047
; Publication No. US20030236392A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20030236392A1e1 full length cDNA
; FILE REFERENCE: HI-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2492
; LENGTH: 244
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-104-047-2492

Query Match          79.3%; Score 23; DB 15; Length 244;
Best Local Similarity 71.4%; Pred. No. 1.1e+03;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy      1 VMXVAEF 7
Db      164 VMVAKF 170

RESULT 67
US-10-334-143-182
; Sequence 182, Application US/10334143
; Publication No. US20040009549A1
; GENERAL INFORMATION:
; APPLICANT: GRIGORIEV, IGOR VYACHESLAVOVICH
; APPLICANT: SUDARSANAM, SUCHA
; TITLE OF INVENTION: METHOD FOR DETECTING REMOTE HOMOLOGUES AND NOVEL
; FILE REFERENCE: 038602/1543
; CURRENT APPLICATION NUMBER: US/10/334,143
; CURRENT FILING DATE: 2002-12-31
; PRIOR APPLICATION NUMBER: 60/343,169
```

```
; PRIOR FILING DATE: 2001-12-31
; NUMBER OF SEQ ID NOS: 207
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 182
; LENGTH: 245
; TYPE: PRT
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: 1fgkA protein sequence
US-10-334-143-182

Query Match          79.3%; Score 23; DB 15; Length 245;
Best Local Similarity 57.1%; Pred. No. 1.1e+03;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
   || :||
Db 131 VMKIADF 137

RESULT 68
US-10-732-923-15746
; Sequence 15746, Application US/10732923
; Publication No. US20050108791A1
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 15746
; LENGTH: 249
; TYPE: PRT
; ORGANISM: Aedes aegypti
US-10-732-923-15746

Query Match          79.3%; Score 23; DB 17; Length 249;
Best Local Similarity 83.3%; Pred. No. 1.2e+03;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 MXVAEF 7
   |||||
Db 25 MLVAEF 30

RESULT 69
US-10-732-923-15615
; Sequence 15615, Application US/10732923
; Publication No. US20050108791A1
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 15615
; LENGTH: 251
; TYPE: PRT
; ORGANISM: Anopheles gambiae str. PEST
US-10-732-923-15615

Query Match          79.3%; Score 23; DB 17; Length 251;
Best Local Similarity 83.3%; Pred. No. 1.2e+03;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 MXVAEF 7
   |||||

; PRIOR FILING DATE: 2001-12-31
; NUMBER OF SEQ ID NOS: 207
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 182
; LENGTH: 245
; TYPE: PRT
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: 1fgkA protein sequence
US-10-334-143-182

Query Match          79.3%; Score 23; DB 15; Length 245;
Best Local Similarity 57.1%; Pred. No. 1.1e+03;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
   || :||
Db 131 VMKIADF 137

RESULT 70
US-10-767-701-42888
; Sequence 42888, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 42888
; LENGTH: 265
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C29394_1.pep
US-10-767-701-42888

Query Match          79.3%; Score 23; DB 16; Length 265;
Best Local Similarity 57.1%; Pred. No. 1.2e+03;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
   || :||
Db 96 VMSIAKF 102

RESULT 71
US-10-470-957-23
; Sequence 23, Application US/10470957
; Publication No. US20040142404A1
; GENERAL INFORMATION:
; APPLICANT: Cytopia Pty Ltd
; APPLICANT: Wilks, Andrew Frederick
; APPLICANT: Atkin, Julie
; APPLICANT: Fantino, Emmanuelle
; TITLE OF INVENTION: Protein kinase signalling
; FILE REFERENCE: 529282001000
; CURRENT APPLICATION NUMBER: US/10/470,957
; CURRENT FILING DATE: 2003-07-30
; PRIOR APPLICATION NUMBER: PCT/AU02/00088
; PRIOR FILING DATE: 2002-01-30
; PRIOR APPLICATION NUMBER: PR 2791
; PRIOR FILING DATE: 2001-01-30
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 23
; LENGTH: 274
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-470-957-23

Query Match          79.3%; Score 23; DB 16; Length 274;
Best Local Similarity 57.1%; Pred. No. 1.3e+03;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
   || :||
Db 150 VMKIADF 156

RESULT 72
US-09-738-626-4505
; Sequence 4505, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
```

```
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: Patent in ver. 3.0
; SEQ ID NO 4505
; LENGTH: 280
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
; OTHER INFORMATION: Description of Unknown Organism: FGFR1 tyrosine kinase
US-09-738-626-4505

Query Match          79.3%; Score 23; DB 9; Length 280;
Best Local Similarity 71.4%; Pred. No. 1.3e+03;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
Db 6 VVVVAEF 12

RESULT 73
US-10-424-599-147161
; Sequence 147161, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 147161
; LENGTH: 287
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_103906C.1.pap
US-10-424-599-147161

Query Match          79.3%; Score 23; DB 15; Length 287;
Best Local Similarity 83.3%; Pred. No. 1.3e+03;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 MXVAEF 7
Db 162 MNVAEF 167

RESULT 74
US-10-377-268-29
; Sequence 29, Application US/10377268
; Publication No. US20040171062A1
```

```
; GENERAL INFORMATION:
; APPLICANT: HIRTH, KLAUS-PETER
; APPLICANT: MILBURN, MICHAEL VANCE
; TITLE OF INVENTION: METHOD FOR THE DESIGN OF MOLECULAR SCAFFOLDS AND LIGANDS
; FILE REFERENCE: 039363/0303
; CURRENT APPLICATION NUMBER: US/10/377,268
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: 60/437,929
; PRIOR FILING DATE: 2003-01-02
; PRIOR APPLICATION NUMBER: 60/360,651
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/411,398
; PRIOR FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: 60/412,341
; PRIOR FILING DATE: 2002-09-20
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 29
; LENGTH: 290
; TYPE: PRT
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: FGFR1 tyrosine kinase
US-10-377-268-29

Query Match          79.3%; Score 23; DB 16; Length 290;
Best Local Similarity 57.1%; Pred. No. 1.4e+03;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
Db 165 VMKIADF 171

RESULT 75
US-10-282-122A-52955
; Sequence 52955, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
```

```
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 52955
; LENGTH: 293
; TYPE: PRT
; ORGANISM: Clostridium difficile
; US-10-282-122A-52955

Query Match          79.3%; Score 23; DB 15; Length 293;
Best Local Similarity 71.4%; Pred. No. 1.4e+03;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 VMXVAEF 7
Db      107 VMVAKF 113

RESULT 76
US-10-334-143-203
; Sequence 203, Application US/10334143
; Publication No. US20040009549A1
; GENERAL INFORMATION:
; APPLICANT: GRIGORIEV, IGOR VYACHESLAVOVICH
; APPLICANT: SUDARSANAM, SUCHA
; TITLE OF INVENTION: METHOD FOR DETECTING REMOTE HOMOLOGUES AND NOVEL
; TITLE OF INVENTION: KINASES IDENTIFIED WITH THE METHOD
; FILE REFERENCE: 038602/1543
; CURRENT APPLICATION NUMBER: US/10/334,143
; CURRENT FILING DATE: 2002-12-31
; PRIOR APPLICATION NUMBER: 60/343,169
; PRIOR FILING DATE: 2001-12-31
; NUMBER OF SEQ ID NOS: 207
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 203
; LENGTH: 299
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-334-143-203

Query Match          79.3%; Score 23; DB 15; Length 299;
Best Local Similarity 57.1%; Pred. No. 1.4e+03;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      1 VMXVAEF 7
Db      173 VMKIADF 179

RESULT 77
US-10-763-418-13
; Sequence 13, Application US/10763418
; Publication No. US20040185547A1
; GENERAL INFORMATION:
; APPLICANT: MOHAMMADI, MOOSA
; APPLICANT: SCHLESSINGER, JOSEPH
; APPLICANT: HUBBARD, STEVAN R.
; TITLE OF INVENTION: CRYSTALS OF THE TYROSINE KINASE DOMAIN OF NON-INSULIN
; TITLE OF INVENTION: RECEPTOR TYROSINE KINASES
; FILE REFERENCE: 038602/0847
; CURRENT APPLICATION NUMBER: US/10/763,418
; CURRENT FILING DATE: 2004-01-26
; PRIOR APPLICATION NUMBER: US/09/664,526
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: 09/188,809
; PRIOR FILING DATE: 1998-11-09
; PRIOR APPLICATION NUMBER: 08/701,191
; PRIOR FILING DATE: 1996-08-21
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.1
; US-10-763-418-13
```

```
; SEQ ID NO 13
; LENGTH: 299
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-763-418-13

Query Match          79.3%; Score 23; DB 16; Length 299;
Best Local Similarity 57.1%; Pred. No. 1.4e+03;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      1 VMXVAEF 7
Db      172 VMKIADF 178

RESULT 78
US-10-763-418-31
; Sequence 31, Application US/10763418
; Publication No. US20040185547A1
; GENERAL INFORMATION:
; APPLICANT: MOHAMMADI, MOOSA
; APPLICANT: SCHLESSINGER, JOSEPH
; APPLICANT: HUBBARD, STEVAN R.
; TITLE OF INVENTION: CRYSTALS OF THE TYROSINE KINASE DOMAIN OF NON-INSULIN
; TITLE OF INVENTION: RECEPTOR TYROSINE KINASES
; FILE REFERENCE: 038602/0847
; CURRENT APPLICATION NUMBER: US/10/763,418
; CURRENT FILING DATE: 2004-01-26
; PRIOR APPLICATION NUMBER: US/09/664,526
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: 09/188,809
; PRIOR FILING DATE: 1998-11-09
; PRIOR APPLICATION NUMBER: 08/701,191
; PRIOR FILING DATE: 1996-08-21
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 31
; LENGTH: 300
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-763-418-31

Query Match          79.3%; Score 23; DB 16; Length 300;
Best Local Similarity 57.1%; Pred. No. 1.4e+03;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      1 VMXVAEF 7
Db      172 VMKIADF 178

RESULT 79
US-10-470-957-21
; Sequence 21, Application US/10470957
; Publication No. US20040142404A1
; GENERAL INFORMATION:
; APPLICANT: Cytopia Pty Ltd
; APPLICANT: Wilks, Andrew Frederick
; APPLICANT: Atkin, Julie
; APPLICANT: Fantino, Emmanuelle
; TITLE OF INVENTION: Protein kinase signalling
; FILE REFERENCE: 529282001000
; CURRENT APPLICATION NUMBER: US/10/470,957
; CURRENT FILING DATE: 2003-07-30
; PRIOR APPLICATION NUMBER: PCT/AU02/00088
; PRIOR FILING DATE: 2002-01-30
; PRIOR APPLICATION NUMBER: PR 2791
; PRIOR FILING DATE: 2001-01-30
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 21
; LENGTH: 301
; TYPE: PRT
; US-10-470-957-21
```



```
; ORGANISM: homo sapiens
US-10-470-957-21

Query Match      79.3%; Score 23; DB 16; Length 301;
Best Local Similarity 57.1%; Pred. No. 1.4e+03;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      1 VMXVAEF 7
Db      176 VKIADF 182

RESULT 80
US-10-369-493-11118
; Sequence 9, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 11118
; LENGTH: 309
; TYPE: PRT
; ORGANISM: Methanosarcina mazei
US-10-369-493-11118

Query Match      79.3%; Score 23; DB 15; Length 309;
Best Local Similarity 83.3%; Pred. No. 1.5e+03;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 MXVAEF 7
Db      264 MKVAEF 269

RESULT 81
US-10-763-418-9
; Sequence 9, Application US/10763418
; Publication No. US20040185547A1
; GENERAL INFORMATION:
; APPLICANT: MOHAMMADI, MOOSA
; APPLICANT: SCHLESSINGER, JOSEPH
; APPLICANT: HUBBARD, STEVAN R.
; TITLE OF INVENTION: CRYSTALS OF THE TYROSINE KINASE DOMAIN OF NON-INSULIN
; FILE OF INVENTION: RECEPTOR TYROSINE KINASES
; FILE REFERENCE: 038602/0847
; CURRENT APPLICATION NUMBER: US/10/763,418
; CURRENT FILING DATE: 2004-01-26
; PRIOR APPLICATION NUMBER: US/09/664,526
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: 09/188,809
; PRIOR FILING DATE: 1998-11-09
; PRIOR APPLICATION NUMBER: 08/701,191
; PRIOR FILING DATE: 1996-08-21
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 309
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-763-418-9

Query Match      79.3%; Score 23; DB 16; Length 309;

; ORGANISM: homo sapiens
US-10-470-957-21

Query Match      79.3%; Score 23; DB 16; Length 301;
Best Local Similarity 57.1%; Pred. No. 1.4e+03;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      1 VMXVAEF 7
Db      176 VKIADF 182

RESULT 82
US-09-939-754-7
; Sequence 7, Application US/09939754
; Patent No. US20020051965A1
; GENERAL INFORMATION:
; APPLICANT: MCTIGUE, MICHELE A.
; APPLICANT: WICKERSHAM, JOHN A.
; APPLICANT: PINKO, CHRIS
; APPLICANT: SHOWALTER, RICHARD
; APPLICANT: PARAST, CAMRAN V.
; APPLICANT: TEMPZYK-RUSSEL, ANNA
; APPLICANT: GEHRING, MICHAEL R.
; APPLICANT: MROCKOWSKI, BARBARA
; APPLICANT: KAN, CHEN-CHEN
; APPLICANT: VILLAFRANCA, J. ERNEST
; APPLICANT: APPELT, KRZYSZTOF
; TITLE OF INVENTION: MODIFICATIONS OF THE VEGF RECEPTOR-2 PROTEIN AND
; FILE OF INVENTION: METHODS OF USE
; FILE REFERENCE: 0125-0016US
; CURRENT APPLICATION NUMBER: US/09/939,754
; CURRENT FILING DATE: 2001-08-28
; PRIOR APPLICATION NUMBER: 09/390,326
; PRIOR FILING DATE: 1999-09-07
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 310
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-939-754-7

Query Match      79.3%; Score 23; DB 9; Length 310;
Best Local Similarity 57.1%; Pred. No. 1.5e+03;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      1 VMXVAEF 7
Db      181 VKIADF 187

RESULT 83
US-09-939-832-7
; Sequence 7, Application US/09939832
; Patent No. US20020127538A1
; GENERAL INFORMATION:
; APPLICANT: MCTIGUE, MICHELE A.
; APPLICANT: WICKERSHAM, JOHN A.
; APPLICANT: PINKO, CHRIS
; APPLICANT: SHOWALTER, RICHARD
; APPLICANT: PARAST, CAMRAN V.
; APPLICANT: TEMPZYK-RUSSEL, ANNA
; APPLICANT: GEHRING, MICHAEL R.
; APPLICANT: MROCKOWSKI, BARBARA
; APPLICANT: KAN, CHEN-CHEN
; APPLICANT: VILLAFRANCA, J. ERNEST
; APPLICANT: APPELT, KRZYSZTOF
; TITLE OF INVENTION: MODIFICATIONS OF THE VEGF RECEPTOR-2 PROTEIN AND
; FILE OF INVENTION: METHODS OF USE
; FILE REFERENCE: 0125-0016US
; CURRENT APPLICATION NUMBER: US/09/939,832
; CURRENT FILING DATE: 2001-08-28
; PRIOR APPLICATION NUMBER: 09/390,326
; PRIOR FILING DATE: 1999-09-07
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
```

```
; SEQ ID NO 7
; LENGTH: 310
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-939-832-7

Query Match          79.3%; Score 23; DB 9; Length 310;
Best Local Similarity 57.1%; Pred. No. 1.5e+03;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VMXVAEF 7
Db 181 VMKIADF 187

RESULT 84
US-09-939-833-7
; Sequence 7, Application US/09939833
; Patent No. US20020164641A1
; GENERAL INFORMATION:
; APPLICANT: MCTIGUE, MICHELE A.
; APPLICANT: WICKERSHAM, JOHN A.
; APPLICANT: PINKO, CHRIS
; APPLICANT: SHOWALTER, RICHARD
; APPLICANT: PARAST, CAMRAN V.
; APPLICANT: TEMPCTZK-RUSSEL, ANNA
; APPLICANT: GEHRING, MICHAEL R.
; APPLICANT: MROCKZKOWSKI, BARBARA
; APPLICANT: KAN, CHEN-CHEN
; APPLICANT: VILLAFRANCA, J. ERNEST
; APPLICANT: APELT, KRZYSZTOF
; TITLE OF INVENTION: MODIFICATIONS OF THE VEGF RECEPTOR-2 PROTEIN AND
; TITLE OF INVENTION: METHODS OF USE
; FILE REFERENCE: 0125-0016US
; CURRENT APPLICATION NUMBER: US/09/939,833
; CURRENT FILING DATE: 2001-08-28
; PRIOR APPLICATION NUMBER: US/09/390,326
; PRIOR FILING DATE: 1999-09-07
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 310
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-939-833-7

Query Match          79.3%; Score 23; DB 9; Length 310;
Best Local Similarity 57.1%; Pred. No. 1.5e+03;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VMXVAEF 7
Db 181 VMKIADF 187

RESULT 85
US-10-664-421-27
; Sequence 27, Application US/10664421
; Publication No. US20040142864A1
; GENERAL INFORMATION:
; APPLICANT: BREWER, RYAN
; APPLICANT: IBRAHIM, PRABHA
; APPLICANT: KUMAR, ABHINAV
; APPLICANT: MANDIYAN, VALSAN
; APPLICANT: MILBURN, MICHAEL V.
; TITLE OF INVENTION: CRYSTAL STRUCTURE OF PIM-1 KINASE
; FILE REFERENCE: 039363/0703
; CURRENT APPLICATION NUMBER: US/10/664,421
; CURRENT FILING DATE: 2003-09-16
; PRIOR APPLICATION NUMBER: 60/412,341
; PRIOR FILING DATE: 2002-09-20
; PRIOR APPLICATION NUMBER: 60/411,398
; PRIOR FILING DATE: 2002-09-16
```

```
; NUMBER OF SEQ ID NOS: 169
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 27
; LENGTH: 310
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-664-421-27

Query Match          79.3%; Score 23; DB 16; Length 310;
Best Local Similarity 57.1%; Pred. No. 1.5e+03;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VMXVAEF 7
Db 181 VMKIADF 187

RESULT 86
US-10-763-418-1
; Sequence 1, Application US/10763418
; Publication No. US20040185547A1
; GENERAL INFORMATION:
; APPLICANT: MOHAMMADI, MOOSA
; APPLICANT: SCHLESSINGER, JOSEPH
; APPLICANT: HUBBARD, STEVAN R.
; TITLE OF INVENTION: CRYSTALS OF THE TYROSINE KINASE DOMAIN OF NON-INSULIN
; FILE REFERENCE: 038602/0847
; CURRENT APPLICATION NUMBER: US/10/763,418
; CURRENT FILING DATE: 2004-01-26
; PRIOR APPLICATION NUMBER: US/09/664,526
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: 09/188,809
; PRIOR FILING DATE: 1998-11-09
; PRIOR APPLICATION NUMBER: 08/701,191
; PRIOR FILING DATE: 1996-08-21
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 310
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-763-418-1

Query Match          79.3%; Score 23; DB 16; Length 310;
Best Local Similarity 57.1%; Pred. No. 1.5e+03;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VMXVAEF 7
Db 181 VMKIADF 187

RESULT 87
US-10-763-418-6
; Sequence 6, Application US/10763418
; Publication No. US20040185547A1
; GENERAL INFORMATION:
; APPLICANT: MOHAMMADI, MOOSA
; APPLICANT: SCHLESSINGER, JOSEPH
; APPLICANT: HUBBARD, STEVAN R.
; TITLE OF INVENTION: CRYSTALS OF THE TYROSINE KINASE DOMAIN OF NON-INSULIN
; FILE REFERENCE: 038602/0847
; CURRENT APPLICATION NUMBER: US/10/763,418
; CURRENT FILING DATE: 2004-01-26
; PRIOR APPLICATION NUMBER: US/09/664,526
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: 09/188,809
; PRIOR FILING DATE: 1998-11-09
; PRIOR APPLICATION NUMBER: 08/701,191
; PRIOR FILING DATE: 1996-08-21
; NUMBER OF SEQ ID NOS: 41
```

```
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 310
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-763-418-6

Query Match          79.3%; Score 23; DB 16; Length 310;
Best Local Similarity 57.1%; Pred. No. 1.5e+03;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      1 VMXVAEF 7
Db      181 VMKIADF 187

RESULT 88
US-10-763-418-7
; Sequence 7, Application US/10763418
; Publication No. US20040185547A1
; GENERAL INFORMATION:
; APPLICANT: MOHAMMADI, MOOSA
; APPLICANT: SCHLESSINGER, JOSEPH
; APPLICANT: HUBBARD, STEVAN R.
; TITLE OF INVENTION: CRYSTALS OF THE TYROSINE KINASE DOMAIN OF NON-INSULIN
; FILE OF INVENTION: RECEPTOR TYROSINE KINASES
; FILE REFERENCE: 038602/0847
; CURRENT APPLICATION NUMBER: US/10/763,418
; CURRENT FILING DATE: 2004-01-26
; PRIOR APPLICATION NUMBER: US/09/664,526
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: 09/188,809
; PRIOR FILING DATE: 1998-11-09
; PRIOR APPLICATION NUMBER: 08/701,191
; PRIOR FILING DATE: 1996-08-21
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 310
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-763-418-7

Query Match          79.3%; Score 23; DB 16; Length 310;
Best Local Similarity 57.1%; Pred. No. 1.5e+03;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      1 VMXVAEF 7
Db      181 VMKIADF 187

RESULT 89
US-10-763-418-8
; Sequence 8, Application US/10763418
; Publication No. US20040185547A1
; GENERAL INFORMATION:
; APPLICANT: MOHAMMADI, MOOSA
; APPLICANT: SCHLESSINGER, JOSEPH
; APPLICANT: HUBBARD, STEVAN R.
; TITLE OF INVENTION: CRYSTALS OF THE TYROSINE KINASE DOMAIN OF NON-INSULIN
; FILE OF INVENTION: RECEPTOR TYROSINE KINASES
; FILE REFERENCE: 038602/0847
; CURRENT APPLICATION NUMBER: US/10/763,418
; CURRENT FILING DATE: 2004-01-26
; PRIOR APPLICATION NUMBER: US/09/664,526
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: 09/188,809
; PRIOR FILING DATE: 1998-11-09
; PRIOR APPLICATION NUMBER: 08/701,191
; PRIOR FILING DATE: 1996-08-21
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 310
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-763-418-8

Query Match          79.3%; Score 23; DB 16; Length 310;
Best Local Similarity 57.1%; Pred. No. 1.5e+03;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      1 VMXVAEF 7
Db      181 VMKIADF 187

RESULT 90
US-09-912-976-70
; Sequence 70, Application US/09912976
; Publication No. US20030212255A1
; GENERAL INFORMATION:
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Mezes, Peter
; APPLICANT: Burgess, Catherine
; APPLICANT: Casman, Stacie
; APPLICANT: Grosse, William M
; APPLICANT: Alsobrook II, John P
; APPLICANT: Lepley, Denise M
; APPLICANT: Gerlach, Valerie L
; APPLICANT: MacDougall, John R
; APPLICANT: Smithson, Glennda
; APPLICANT: Mishra, Vishnu
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-070
; CURRENT APPLICATION NUMBER: US/09/912,976
; CURRENT FILING DATE: 2001-07-05
; PRIOR APPLICATION NUMBER: 60/221,336
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/238,333
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/260,675
; PRIOR FILING DATE: 2001-01-10
; PRIOR APPLICATION NUMBER: 60/271,025
; PRIOR FILING DATE: 2001-02-22
; PRIOR APPLICATION NUMBER: 60/278,164
; PRIOR FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: 60/280,876
; PRIOR FILING DATE: 2001-04-02
; NUMBER OF SEQ ID NOS: 99
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 70
; LENGTH: 312
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-912-976-70

Query Match          79.3%; Score 23; DB 10; Length 312;
Best Local Similarity 57.1%; Pred. No. 1.5e+03;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      1 VMXVAEF 7
Db      196 IMGVADF 202

RESULT 91
US-10-774-355A-2526
; Sequence 2526, Application US/10774355A
; Publication No. US20050043513A1
; GENERAL INFORMATION:
; APPLICANT: Firestein, Stuart
; APPLICANT: Zhang, Xinmin
; TITLE OF INVENTION: MOUSE OLFACTORY RECEPTOR GENE SUPERFAMILY
; FILE REFERENCE: A34570-PCT-USA-A 070050.2520
```

```
; CURRENT APPLICATION NUMBER: US/10/774,355A
; CURRENT FILING DATE: 2004-02-06
; PRIOR APPLICATION NUMBER: PCT/US02/25556
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 60/311,159
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 60/339,694
; PRIOR FILING DATE: 2001-12-12
; NUMBER OF SEQ ID NOS: 2596
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2526
; LENGTH: 312
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-774-355A-2526

Query Match          79.3%; Score 23; DB 17; Length 312;
Best Local Similarity 57.1%; Pred. No. 1.5e+03;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      1 VMXVAEF 7
      :|:|:|
Db      196 IMGVADF 202

RESULT 92
US-09-510-332-44
; Sequence 44, Application US/09510332
; Publication No. US20030022278A1
; GENERAL INFORMATION:
; APPLICANT: Zuker, Charles S.
; APPLICANT: Adler, Jon Elliot
; APPLICANT: Ryba, Nick
; APPLICANT: Mueller, Ken
; APPLICANT: Hoon, Mark
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: T2R, a No. US20030022278A1el Family of Taste Receptors
; FILE REFERENCE: 02307E-098010US
; CURRENT APPLICATION NUMBER: US/09/510,332
; CURRENT FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 09/393,634
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 172
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 44
; LENGTH: 314
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: human T2R24 (hGR24)
US-09-510-332-44

Query Match          79.3%; Score 23; DB 10; Length 314;
Best Local Similarity 42.9%; Pred. No. 1.5e+03;
Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY      1 VMXVAEF 7
      :|:|:|
Db      11 ILAIAEF 17

RESULT 93
US-09-510-332-71
; Sequence 71, Application US/09510332
; Publication No. US20030022278A1
; GENERAL INFORMATION:
; APPLICANT: Zuker, Charles S.
; APPLICANT: Adler, Jon Elliot
; APPLICANT: Ryba, Nick
; APPLICANT: Mueller, Ken
; APPLICANT: Hoon, Mark
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: T2R, a No. US20030022278A1el Family of Taste Receptors
; FILE REFERENCE: 02307E-098010US
; CURRENT APPLICATION NUMBER: US/09/510,332
; CURRENT FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 09/393,634
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 172
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 44
; LENGTH: 314
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: human T2R24 (hGR24)
US-09-510-332-71

Query Match          79.3%; Score 23; DB 10; Length 314;
Best Local Similarity 42.9%; Pred. No. 1.5e+03;
Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY      1 VMXVAEF 7
      :|:|:|
Db      11 ILAIAEF 17

RESULT 94
US-10-017-161-1934
; Sequence 1934, Application US/10017161
; Publication No. US20030143668A1
; GENERAL INFORMATION:
; APPLICANT: SUWA, MAKIKO
; APPLICANT: ASAI, KIYOSHI
; APPLICANT: AKIYAMA, YUTAKA
; APPLICANT: ABURATANI, HIROYUKI
; TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTORS
; FILE REFERENCE: 084335/0152
; CURRENT APPLICATION NUMBER: US/10/017,161
; CURRENT FILING DATE: 2002-12-18
; PRIOR APPLICATION NUMBER: JP 2001/246789
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 2430
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1934
; LENGTH: 314
; TYPE: PRT
; ORGANISM: Homo sapiens
; OTHER INFORMATION: human T2R45 (hGR45)
US-10-017-161-1934

Query Match          79.3%; Score 23; DB 10; Length 314;
Best Local Similarity 42.9%; Pred. No. 1.5e+03;
Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY      1 VMXVAEF 7
      :|:|:|
Db      11 ILAIAEF 17

RESULT 95
US-10-292-798-1478
; Sequence 1478, Application US/10292798
; Publication No. US20030235833A1
; GENERAL INFORMATION:
; APPLICANT: SUWA, MAKIKO
; APPLICANT: ASAI, KIYOSHI
; APPLICANT: AKIYAMA, YUTAKA
; APPLICANT: ABURATANI, HIROYUKI
; TITLE OF INVENTION: GUANOSINE TRIPHOSPHATE-BINDING PROTEIN COUPLED RECEPTORS
; FILE REFERENCE: 084335/166
; CURRENT APPLICATION NUMBER: US/10/292,798
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: 10/017,161
; PRIOR FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: JP 2001-246789
; PRIOR FILING DATE: 2001-06-18
```

```
; NUMBER OF SEQ ID NOS: 2070
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1478
; LENGTH: 314
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-292-798-1478

Query Match          79.3%; Score 23; DB 15; Length 314;
Best Local Similarity 42.9%; Pred. No. 1.5e+03;
Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
Db 11 ILAIAEF 17

RESULT 96
US-10-343-650A-672
; Sequence 672, Application US/10343650A
; Publication No. US20040067499A1
; GENERAL INFORMATION:
; APPLICANT: HAGA, TATSUYA
; TITLE OF INVENTION: NOVEL G-PROTEIN COUPLED RECEPTOR
; FILE REFERENCE: 31671-186347
; CURRENT APPLICATION NUMBER: US/10/343,650A
; CURRENT FILING DATE: 2003-07-21
; PRIOR APPLICATION NUMBER: JP 2000/237818
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: JP 2001/34434
; PRIOR FILING DATE: 2001-02-13
; NUMBER OF SEQ ID NOS: 694
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 672
; LENGTH: 314
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-343-650A-672

Query Match          79.3%; Score 23; DB 15; Length 314;
Best Local Similarity 42.9%; Pred. No. 1.5e+03;
Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
Db 11 ILAIAEF 17

RESULT 97
US-10-770-127-44
; Sequence 44, Application US/10770127
; Publication No. US20040214239A1
; GENERAL INFORMATION:
; APPLICANT: SERVANT, GUY
; APPLICANT: OZECK, MARK
; APPLICANT: BRUST, PAUL
; APPLICANT: XU, HONG
; TITLE OF INVENTION: FUNCTIONAL COUPLING OF T1RS AND T2RS BY GI PROTEINS
; TITLE OF INVENTION: AND CELL-BASED ASSAYS FOR THE IDENTIFICATION OF T1R
; FILE REFERENCE: 100337.54281US
; CURRENT APPLICATION NUMBER: US/10/770,127
; CURRENT FILING DATE: 2004-02-03
; PRIOR APPLICATION NUMBER: 60/444,172
; PRIOR FILING DATE: 2003-02-03
; PRIOR APPLICATION NUMBER: 60/457,318
; PRIOR FILING DATE: 2003-03-26
; NUMBER OF SEQ ID NOS: 210
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 44
; LENGTH: 314
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-770-127-44

Query Match          79.3%; Score 23; DB 16; Length 314;
Best Local Similarity 42.9%; Pred. No. 1.5e+03;
Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
Db 11 ILAIAEF 17

RESULT 98
US-10-770-127-71
; Sequence 71, Application US/10770127
; Publication No. US20040214239A1
; GENERAL INFORMATION:
; APPLICANT: SERVANT, GUY
; APPLICANT: OZECK, MARK
; APPLICANT: BRUST, PAUL
; APPLICANT: XU, HONG
; TITLE OF INVENTION: FUNCTIONAL COUPLING OF T1RS AND T2RS BY GI PROTEINS
; TITLE OF INVENTION: AND CELL-BASED ASSAYS FOR THE IDENTIFICATION OF T1R
; FILE REFERENCE: 100337.54281US
; CURRENT APPLICATION NUMBER: US/10/770,127
; CURRENT FILING DATE: 2004-02-03
; PRIOR APPLICATION NUMBER: 60/444,172
; PRIOR FILING DATE: 2003-02-03
; PRIOR APPLICATION NUMBER: 60/457,318
; PRIOR FILING DATE: 2003-03-26
; NUMBER OF SEQ ID NOS: 210
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 71
; LENGTH: 314
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-770-127-71

Query Match          79.3%; Score 23; DB 16; Length 314;
Best Local Similarity 42.9%; Pred. No. 1.5e+03;
Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
Db 11 ILAIAEF 17

RESULT 99
US-10-962-365-44
; Sequence 44, Application US/10962365
; Publication No. US20050048586A1
; GENERAL INFORMATION:
; APPLICANT: ZUKER, Charles S.
; APPLICANT: Adler, Jon Elliot
; APPLICANT: Ryba, Nick
; APPLICANT: Mueller, Ken
; APPLICANT: Hoon, Mark
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: T2R, a Novel Family of Taste Receptors
; FILE REFERENCE: 02307E-098010US
; CURRENT APPLICATION NUMBER: US/10/962,365
; CURRENT FILING DATE: 2004-10-07
; PRIOR APPLICATION NUMBER: US/09/510,332
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 09/393,634
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 172
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 44
; LENGTH: 314
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-962-365-44
```

```
; OTHER INFORMATION: human T2R24 (hGR24)
US-10-962-365-44

Query Match          79.3%; Score 23; DB 17; Length 314;
Best Local Similarity 42.9%; Pred. No. 1.5e+03;
Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy      1 VMXVAEF 7
Db      11 ILAIAEF 17

RESULT 100
US-10-962-365-71
; Sequence 71, Application US/10962365
; Publication No. US20050048586A1
; GENERAL INFORMATION:
; APPLICANT: Zuker, Charles S.
; APPLICANT: Adler, Jon Elliot
; APPLICANT: Ryba, Nick
; APPLICANT: Mueller, Ken
; APPLICANT: Hoon, Mark
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: T2R, a Novel Family of Taste Receptors
; FILE REFERENCE: 02307E-098010US
; CURRENT APPLICATION NUMBER: US/10/962,365
; CURRENT FILING DATE: 2004-10-07
; PRIOR APPLICATION NUMBER: US/09/510,332
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 09/393,634
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 172
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 71
; LENGTH: 314
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: human T2R45 (hGR45)
US-10-962-365-71

Query Match          79.3%; Score 23; DB 17; Length 314;
Best Local Similarity 42.9%; Pred. No. 1.5e+03;
Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy      1 VMXVAEF 7
Db      11 ILAIAEF 17

Search completed: June 13, 2005, 14:18:32
Job time : 117 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2005, 13:36:32 ; Search time 24 Seconds
(without alignments)
28.063 Million cell updates/sec

Title: 09730329-60ED

Perfect score: 29

Sequence: 1 VMXVAEF 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 300 summaries

Database :

PIR_79:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	28	96.6	2638	1	genome polyprotein
2	27	93.1	802	2	probable potaesium
3	26	89.7	223	2	deoxyribose-phosph
4	26	89.7	224	2	deoxyribose-phosph
5	26	89.7	324	2	probable steroid s
6	25	86.2	201	2	hypothetical prote
7	25	86.2	632	2	probable electron
8	24	82.8	145	1	ribosomal protein
9	24	82.8	135	2	thymidine kinase (
10	24	82.8	236	2	hypothetical prote
11	24	82.8	248	2	hypothetical prote
12	24	82.8	249	2	serine/threonine p
13	24	82.8	335	2	6-phosphofructokin
14	24	82.8	335	2	6-phosphofructokin
15	24	82.8	367	2	fusion, chorismate
16	24	82.8	484	2	unknown protein, 3
17	24	82.8	565	2	probable type II D
18	24	82.8	582	2	hypothetical prote
19	24	82.8	584	2	archaeosine trna-r
20	24	82.8	754	2	hypothetical prote
21	24	82.8	919	2	hypothetical prote
22	24	82.8	921	2	isoleucine-trna li
23	24	82.8	921	2	isoleucyl-trna syn
24	24	82.8	1122	2	l42400
25	23	79.3	89	2	Eph receptor tyros
26	23	79.3	96	2	hypothetical prote
27	23	79.3	112	2	hypothetical prote
28	23	79.3	112	2	hypothetical prote
29	23	79.3	143	2	riboflavin synthas

RGF-receptor - com
shikimate kinase (acetyltransferase
fibroblast growth
fibroblast growth
fibroblast growth
photosynthetic rea
transcription regu
probable tonB tran
phosphomannomutase
NADH2 dehydrogenas
methionyl aminopep
C14 protein - rabb
ribonuclease III (genome polyprotein
conserved hypothe
fibroblast growth
probable carbon-su
pantothenate metab
heparin-binding fi
hypothetical prote
fAD/FMN-containing
hypothetical prote
probable aldehyde d
hypothetical prote
arginyl-tRNA synth
protein-tyrosine k
hypothetical prote
fibroblast growth
DNA topoisomerase
DNA gyrase chain B
two-component sens
heparin-binding gr
fibroblast growth
keratinocyte growt
fibroblast growth
fibroblast growth
fibroblast growth
fibroblast growth
DNA gyrase chain B
DNA gyrase, sub B
protein-tyrosine k
hypothetical prote
fibroblast growth
fibroblast growth
heparin-binding gr
fibroblast growth
fibroblast growth
fibroblast growth
protein-tyrosine k
fibroblast growth
fibroblast growth
fibroblast growth
fibroblast growth
fibroblast growth
fibroblast growth
keratinocyte growt
fibroblast growth

103	23	79.3	822	2	B49151	fibroblast growth	176	22	75.9	783	2	T33870	hypothetical prote
104	23	79.3	822	2	I49289	fibroblast growth	177	22	75.9	790	2	T20312	hypothetical prote
105	23	79.3	822	2	S29840	fibroblast growth	178	22	75.9	812	1	A36477	phenylalanyl-tRNA
106	23	79.3	823	2	B35963	protein-tyrosine k	179	22	75.9	820	2	C75284	anion transporter
107	23	79.3	824	2	S24108	protein-tyrosine k	180	22	75.9	844	2	I50159	DNA topoisomerase
108	23	79.3	829	2	JC4583	fibroblast growth	181	22	75.9	902	2	AF3362	band 3 anion trans
109	23	79.3	832	2	JH0393	fibroblast growth	182	22	75.9	922	2	A30816	protein-tyrosine k
110	23	79.3	852	1	JVBEIV	DNA-packaging prot	183	22	75.9	952	2	I50612	receptor protein-t
111	23	79.3	921	2	AC1327	isoleucyl-tRNA syn	184	22	75.9	970	2	I78842	protein-tyrosine k
112	23	79.3	921	2	AE1698	isoleucyl-tRNA syn	185	22	75.9	984	2	A39753	receptor tyrosine k
113	23	79.3	980	2	A38523	genome polyprotein	186	22	75.9	985	2	I51672	receptor tyrosine
114	23	79.3	1052	2	B49120	protein-tyrosine k	187	22	75.9	988	2	I50611	protein-tyrosine k
115	23	79.3	1085	2	G89056	protein K09H11.3 l	188	22	75.9	995	2	A56599	embryo kinase 5 -
116	23	79.3	1256	2	AB2042	hypothetical prote	189	22	75.9	998	2	I58351	receptor protein-t
117	23	79.3	3412	1	GNWVTE	genome polyprotein	190	22	75.9	998	2	JC5672	receptor tyrosine
118	23	79.3	3414	1	GNWVTE	genome polyprotein	191	22	75.9	1188	2	T19552	hypothetical prote
119	23	79.3	3415	2	A46105	polyprotein(NS1, N	192	22	75.9	1259	2	T32901	hypothetical prote
120	22	75.9	120	2	AI0923	conserved hypotet	193	22	75.9	1477	2	T18534	protein-tyrosine k
121	22	75.9	153	2	F96725	hypothetical prote	194	22	75.9	1560	2	T30282	calcium-binding pr
122	22	75.9	155	2	T45215	transcription regu	195	22	75.9	1788	2	T31095	vitellinogen precu
123	22	75.9	170	2	B69204	conserved hypotet	196	22	75.9	2140	2	F95074	serine proteinase,
124	22	75.9	171	2	C84141	hypothetical prote	197	22	75.9	2144	2	A97942	metalloproteinase
125	22	75.9	181	2	T11902	NADH2 dehydrogenas	198	22	75.9	2344	2	T41590	probable sensor-li
126	22	75.9	184	2	C69133	DNA-dependent RNA	199	22	75.9	2361	2	T25752	hypothetical prote
127	22	75.9	185	2	A39658	polyposis coli reg	200	22	75.9	2584	2	T24158	hypothetical prote
128	22	75.9	185	2	JC4667	TB2/DPI protein ho	201	22	75.9	2606	2	T24157	hypothetical prote
129	22	75.9	188	2	B87329	conserved hypotet	202	22	75.9	2693	2	A40743	IP3 receptor, XIPI
130	22	75.9	236	2	A53853	apolipoprotein B m	203	22	75.9	4450	2	JX0340	gramicidin S synth
131	22	75.9	243	1	D70456	guFA protein homol	204	22	75.9	4452	1	YGBSG2	gramicidin S synth
132	22	75.9	254	2	E75052	guFA protein homol	205	22	75.9	5037	1	A54161	ryanodine-binding
133	22	75.9	265	2	B75161	phospho-2-dehydro-	206	21	72.4	46	2	F95318	conserved hypotet
134	22	75.9	269	1	H71107	guFA protein homol	207	21	72.4	71	2	AG2798	hypothetical prote
135	22	75.9	275	2	AC1247	B. subtilis SpoIII	208	21	72.4	77	2	B95003	hypothetical prote
136	22	75.9	275	2	AG1609	B. subtilis SpoIII	209	21	72.4	87	2	B64577	conserved hypotet
137	22	75.9	298	2	T20841	hypothetical prote	210	21	72.4	97	2	E82462	hypothetical prote
138	22	75.9	310	2	A71374	probable hypoteti	211	21	72.4	108	2	C64490	hypothetical prote
139	22	75.9	311	2	T23438	hypothetical prote	212	21	72.4	110	2	A95296	hypothetical prote
140	22	75.9	339	2	D97602	hypothetical prote	213	21	72.4	112	2	F97875	hypothetical prote
141	22	75.9	339	2	AE2824	ketol-acid reducto	214	21	72.4	120	1	D71023	hypothetical prote
142	22	75.9	339	2	G87511	ketol-acid reducto	215	21	72.4	122	2	A24753	ribosomal protein
143	22	75.9	373	2	D71094	probable cofactor	216	21	72.4	125	2	B81993	probable IS1016 tr
144	22	75.9	373	2	D75071	tungsten-contains	217	21	72.4	145	2	B84269	50S ribosomal prot
145	22	75.9	393	2	E95261	serine proteinase	218	21	72.4	145	2	S77043	hypothetical prote
146	22	75.9	397	2	B98127	serine proteinase	219	21	72.4	145	2	A95256	conserved hypotet
147	22	75.9	399	2	H87342	glycosyl transfera	220	21	72.4	150	2	B55209	H transfer determi
148	22	75.9	438	2	E97342	high affinity gluc	221	21	72.4	150	2	A55209	H transfer determi
149	22	75.9	440	2	B41860	tetracycline resis	222	21	72.4	151	2	F82011	probable transposo
150	22	75.9	446	2	S59646	clathrin coat asse	223	21	72.4	151	2	D82030	probable IS1016 tr
151	22	75.9	449	2	S41950	UTP-glucose glucos	224	21	72.4	152	2	A64610	hypothetical prote
152	22	75.9	478	2	A57174	protein-tyrosine k	225	21	72.4	152	2	A71904	hypothetical prote
153	22	75.9	480	2	I48760	protein-tyrosine k	226	21	72.4	168	2	S48866	hypothetical 19.5K
154	22	75.9	488	2	AF2386	neopullulanase [m	227	21	72.4	171	2	D95138	galactose-6-phosph
155	22	75.9	514	2	T10559	hypothetical prote	228	21	72.4	171	2	B39778	galactose-6-phosph
156	22	75.9	548	2	D69187	probable acid-CoA	229	21	72.4	171	2	B98006	galactose-6-phosph
157	22	75.9	573	2	B70047	two-component sens	230	21	72.4	180	2	B83758	chromate transport
158	22	75.9	580	2	S63391	probable membrane	231	21	72.4	180	2	C85358	hypothetical prote
159	22	75.9	593	2	B69378	probable acid-CoA	232	21	72.4	185	2	A64159	intracellular sept
160	22	75.9	598	2	B71095	hypothetical prote	233	21	72.4	188	2	T47912	hypothetical prote
161	22	75.9	601	2	D96001	probable 1,4-alpha	234	21	72.4	191	2	A87365	transcription regu
162	22	75.9	602	1	S38111	amino acid transpo	235	21	72.4	194	1	B9E8R5	resolvase - Escher
163	22	75.9	611	2	T22855	hypothetical prote	236	21	72.4	195	2	D84305	thymidine kinase (
164	22	75.9	613	2	T16885	hypothetical prote	237	21	72.4	196	2	D64909	probable resolvase
165	22	75.9	633	2	C87274	TrkA family protei	238	21	72.4	196	2	A64888	probable resolvase
166	22	75.9	637	2	S66236	acetylcholinestera	239	21	72.4	201	2	C70170	hypothetical prote
167	22	75.9	664	2	T20420	hypothetical prote	240	21	72.4	204	2	T30109	hypothetical prote
168	22	75.9	668	2	H71312	probable ATP-depen	241	21	72.4	213	2	T47123	thymidine kinase (
169	22	75.9	670	2	D86457	hypothetical prote	242	21	72.4	213	2	AF2551	hypothetical prote
170	22	75.9	691	1	S76521	hypothetical prote	243	21	72.4	218	2	A64482	hypothetical prote
171	22	75.9	704	2	F81680	conserved hypotet	244	21	72.4	222	2	AB0007	probable DNA repa
172	22	75.9	734	2	T13674	hypothetical prote	245	21	72.4	232	2	E87520	hypothetical prote
173	22	75.9	743	2	T13673	hypothetical prote	246	21	72.4	239	2	T23169	hypothetical prote
174	22	75.9	758	2	F75590	methyl-accepting c	247	21	72.4	239	2	T36382	hypothetical prote
175	22	75.9	777	2	T44900	probable polyribon	248	21	72.4	242	2	F95169	spou RNA methylas

249 72.4 21 T13456
250 72.4 21 T29733
251 72.4 21 T25121
252 72.4 252 E72060
253 72.4 252 E86564
254 72.4 253 E98035
255 72.4 258 E83275
256 72.4 262 E72333
257 72.4 268 E97189
258 72.4 269 E72550
259 72.4 269 E56487
260 72.4 276 E86756
261 72.4 278 E86421
262 72.4 283 E98065
263 72.4 290 E82030
264 72.4 291 E40740
265 72.4 294 E64465
266 72.4 294 E74562
267 72.4 298 E76437
268 72.4 300 E710312
269 72.4 304 E81417
270 72.4 306 E705274
271 72.4 306 E81159
272 72.4 306 E81518
273 72.4 308 E730315
274 72.4 311 E95117
275 72.4 311 E97986
276 72.4 313 E83095
277 72.4 313 E95853
278 72.4 317 E86292
279 72.4 322 E84320
280 72.4 325 E745994
281 72.4 327 E83482
282 72.4 328 E723498
283 72.4 331 E83117
284 72.4 331 E72783
285 72.4 332 E75465
286 72.4 335 E82686
287 72.4 336 E719802
288 72.4 336 E700571
289 72.4 338 E97467
290 72.4 343 E69211
291 72.4 343 E88429
292 72.4 343 E86821
293 72.4 345 E98170
294 72.4 357 E81276
295 72.4 357 E81639
296 72.4 361 E71362
297 72.4 366 E49964
298 72.4 366 E750991
299 72.4 366 E83398
300 72.4 367 E84258

ALIGNMENTS

RESULT 1
A42545
genome polyprotein - Langat virus (strain TP21) (fragment)
N; Contains: nonstructural protein NS1; nonstructural protein NS2a; nonstructural protein NS3
C; Species: Langat virus
C; Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004
C; Accession: A42545; A61409; C61409
R; Tacono-Connors, L.C.; Schmaljohn, C.S.
Virology 188, 875-880, 1992
A; Title: Cloning and sequence analysis of the genes encoding the nonstructural proteins
A; Reference number: A42545; MUID: 92263794; PMID: 1316684
A; Accession: A42545
A; Molecule type: genomic RNA
A; Residues: 1-2638 <IAC>
A; Cross-references: UNIPROT: P29837; GB: 835365; NID: g249315; PID: AAB22165.1; PID: g24931d

R; Guirakhoo, F.; Heinz, P.X.; Mandl, C.W.; Holzmann, H.; Kunz, C.; Gresikova, M.
J. Gen. Virol. 72, 333-338, 1991
A; Title: The relationship between the flaviviruses Skastica and Langat as revealed by monoclonal antibodies
A; Reference number: A61409; MUID: 911212129; PMID: 1847173
A; Accession: A61409
A; Status: not compared with conceptual translation
A; Molecule type: genomic RNA
A; Residues: 319-337 <GUI>
A; Accession: C61409
A; Status: not compared with conceptual translation
A; Molecule type: genomic RNA
A; Residues: 877-994 <GU2>
C; Superfamily: yellow fever virus genome polyprotein
C; Keywords: glycoprotein; nonstructural protein NS1 #status predicted <NS1>
F; 1-352/Product: nonstructural protein NS2a #status predicted <NS2>
F; 353-582/Product: nonstructural protein NS2b #status predicted <NS2>
F; 583-713/Product: nonstructural protein NS3 #status predicted <NS3>
F; 714-1334/Product: nonstructural protein NS3 #status predicted <NS3>
F; 912-919/Region: nucleotide-binding motif A (P-loop)
F; 1335-1493/Product: nonstructural protein NS4a #status predicted <NS4>
F; 1484-1735/Product: nonstructural protein NS4b #status predicted <NS4>
F; 1736-2638/Product: nonstructural protein NS5 #status predicted <NS5>
F; 85,207,223,873,1212,1671,1950/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 96.6%; Score 28; DB 1; Length 2638;
Best Local Similarity 85.7%; Pred. No. 86;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVASEF 7
|||
DB 153 VMTVASEF 159

RESULT 2

T05596
probable potassium transport protein F9D16.110 - Arabidopsis thaliana
C; Species: Arabidopsis thaliana (mouse-ear cress)
C; Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C; Accession: T05596
R; Bevan, M.; Wedler, H.; Wedler, E.; Wambutt, R.; Hoheisel, J.; Mewes, H.W.; Mayer, K.F.
submitted to the Protein Sequence Database, February 1999
A; Reference number: Z15419
A; Accession: T05596
A; Molecule type: DNA
A; Residues: 1-802 <BEV>
A; Cross-references: UNIPROT: Q9FE38; EMBL: AL035394
A; Experimental source: cultivar Columbia; BAC clone F9D16
C; Genetics:
A; Map position: 4
A; Introns: 1/3; 16/3; 102/1; 185/1; 209/1; 296/1; 313/3; 352/1; 437/1
A; Note: F9D16.110
C; Superfamily: barley probable potassium transport protein HAK1
C; Keywords: ion transport

Query Match 93.1%; Score 27; DB 2; Length 802;
Best Local Similarity 71.4%; Pred. No. 48;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVASEF 7
|||
DB 629 VMSIAEF 635

RESULT 3

E64205
deoxyribose-phosphate aldolase (EC 4.1.2.4) - Mycoplasma genitalium
C; Species: Mycoplasma genitalium
C; Date: 17-Nov-1995 #sequence_revision 17-Nov-1995 #text_change 09-Jul-2004
C; Accession: E64205
R; Fraser, C.W.; Gocayne, J.D.; White, O.; Adams, M.D.; Clayton, R.A.; Fleischmann, R.D.;
M.; Fuhrmann, J.; Nguyen, D.; Utterback, T.R.; Saudek, D.M.; Phillips, C.A.; Merrick, J.I.
C.A.; Venter, J.C.
Science 270, 397-403, 1995

A;Title: The minimal gene complement of Mycoplasma genitalium.
A;Reference number: A64200; MUID:96026346; PMID:7569993
A;Accession: B64205
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-223 <TIGR>
A;Cross-references: UNIPROT:P47296; GB:U39684; GB:U43967; NID:G3844650; PIDN:AACT1266.1;
A;Experimental source: strain G-37
C;Genetics:
A;Genetic code: SGC3
C;Superfamily: deoxyribose-phosphate aldolase
C;Keywords: aldehyde-lyase; carbon-carbon lyase

Query Match 89.7%; Score 26; DB 2; Length 223;
Best Local Similarity 71.4%; Pred. No. 25;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
|||
93 VMNIAEF 99

Db

RESULT 4
S02216
deoxyribose-phosphate aldolase (EC 4.1.2.4) - Mycoplasma pneumoniae
C;Species: Mycoplasma pneumoniae
A;Variety: ATCC 29342
C;Date: 18-Oct-1989 #sequence_revision 18-Oct-1989 #text_change 09-Jul-2004
C;Accession: S02216; S73417
R;Loechel, S.; Inamine, J.M.; Hu, P.C.
Nucleic Acids Res. 17, 801, 1989
A;Title: Nucleotide sequence of the deoC gene of Mycoplasma pneumoniae.
A;Reference number: S02216; MUID:89128453; PMID:2452658
A;Accession: S02216
A;Molecule type: DNA
A;Residues: 1-224 <LOE>
A;Cross-references: UNIPROT:P09924; EMBL:X13544; NID:G44480; PIDN:CAA31897.1; PID:G44480
R;Hammelreich, R.; Hilbert, H.; Plagens, H.; Pirkel, E.; Li, B.C.; Herrmann, R.
Nucleic Acids Res. 24, 4420-4449, 1996
A;Title: Complete sequence analysis of the genome of the bacterium Mycoplasma pneumoniae
A;Reference number: S73327; MUID:97105885; PMID:8948633
A;Accession: S73417
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-224 <HIM>
A;Cross-references: EMBL:AE000011; GB:U00089; NID:G1673740; PIDN:AA895739.1; PID:G167374
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1996
C;Genetics:
A;Gene: deoC
A;Genetic code: SGC3
C;Superfamily: deoxyribose-phosphate aldolase
C;Keywords: aldehyde-lyase; carbon-carbon lyase

Query Match 89.7%; Score 26; DB 2; Length 224;
Best Local Similarity 71.4%; Pred. No. 25;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
|||
93 VMNIAEF 99

Db

RESULT 5
B84452
probable steroid sulfotransferase [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C;Accession: B84452
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter,
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: B84452
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-324 <STO>
A;Cross-references: UNIPROT:Q9ZPQ5; GB:AE002093; NID:G4406768; PIDN:AAD20079.1; GSPDB:GN
C;Genetics:
A;Gene: At2g03770
A;Map position: 2
C;Superfamily: alcohol sulfotransferase

Query Match 89.7%; Score 26; DB 2; Length 324;
Best Local Similarity 71.4%; Pred. No. 36;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
|||
233 VMNIAEF 239

Db

RESULT 6
F69988
hypothetical protein ytbQ - Bacillus subtilis
C;Species: Bacillus subtilis
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 15-Oct-1999
C;Accession: F69988
R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bertero
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Choi
A.; Ehrlich, S.D.; Emmerison, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallen
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.;
Kobter, P.; Konningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel,
Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle,
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon,
A;Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seror,
akeuchi, M.; Tanakoshi, A.; Tanaka, T.; Terpsira, P.; Tognoni, A.; Tosato, V.; Uchiyama,
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.
A;Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A;Reference number: A69580; MUID:98044033; PMID:9384377
A;Accession: F69988
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-201 <KUN>
A;Cross-references: GB:Z99119; GB:AL009126; NID:G2635411; PIDN:CAB14996.1; PID:ell85891;
A;Experimental source: strain 168
C;Genetics:
A;Gene: ytbQ

Query Match 86.2%; Score 25; DB 2; Length 201;
Best Local Similarity 57.1%; Pred. No. 42;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
|||
23 IMDIAEF 29

Db

RESULT 7
T38126
probable electron transfer flavoprotein precursor - fission yeast (Schizosaccharomyces p
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: T38126
R;Badcock, K.; Churcher, C.M.; Wood, V.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, May 1997
A;Reference number: Z21772
A;Accession: T38126
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-632 <BAD>

A;Cross-references: UNIPROT:P87111; EMBL:Z95334; PIDN:CAB08598.1; GSPDB:GN00066; SPDB:SH
A;Experimental source: strain 972h-; cosmid c20G8
C;Genetics:
A;Gene: SPDB:SPAC20G8.04c
A;Map position: 1

Query Match 86.2%; Score 25; DB 2; Length 632;
Best Local Similarity 71.4%; Pred. No. 1.3e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
|||:||||
Db 193 VMSLAEF 199

RESULT 8
B41715
ribosomal protein L13 [similarity] - Haloarcula marismortui
C;Species: Haloarcula marismortui
C;Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 09-Jul-2004
C;Accession: B41715
R;Kroemer, W.J.; Arndt, E.
J. Biol. Chem. 266, 24573-24579, 1991
A;Title: Halobacterial S9 operon. Three ribosomal protein genes are cotranscribed with a
bacterium) marismortui.
A;Reference number: A41715; MUID:92105119; PMID:1840597
A;Accession: B41715
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-145 <KRO>
A;Cross-references: UNIPROT:P29198; GB:M76567; NID:gl48775; PIDN:AAA73097.1; PID:gl48777
C;Superfamily: ribosomal protein L13

Query Match 82.8%; Score 24; DB 1; Length 145;
Best Local Similarity 83.3%; Pred. No. 56;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 MXVAEF 7
|||:||||
Db 1 MSVAEF 6

RESULT 9
C97255
thymidine kinase (EC 2.7.1.21) [similarity] - Clostridium acetobutylicum
C;Species: Clostridium acetobutylicum
C;Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 09-Jul-2004
C;Accession: C97255
R;Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,
J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clo
A;Reference number: A96900; MUID:21359325; PMID:21359325
A;Accession: C97255
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-195 <KUR>
A;Cross-references: UNIPROT:Q97F65; GB:AE001437; PIDN:AAK80830.1; PID:gl5025935; GSPDB:C
A;Experimental source: Clostridium acetobutylicum ATCC824
C;Genetics:
A;Gene: CAC2887
C;Superfamily: thymidine kinase
C;Keywords: phosphotransferase

Query Match 82.8%; Score 24; DB 2; Length 195;
Best Local Similarity 57.1%; Pred. No. 75;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
|||:||||
Db 131 LMAIAEF 137

RESULT 10
B70728
hypothetical protein RV2558 - Mycobacterium tuberculosis (strain H37RV)
C;Species: Mycobacterium tuberculosis
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C;Accession: B70728
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.;
Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.;
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A;Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A;Reference number: A70500; MUID:98295987; PMID:9634230
A;Accession: B70728
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-236 <COL>
A;Cross-references: UNIPROT:Q50740; GB:277250; GB:AL123456; NID:g3261617; PIDN:CAB01046.1
A;Experimental source: strain H37RV
C;Genetics:
A;Gene: RV2558

Query Match 82.8%; Score 24; DB 2; Length 236;
Best Local Similarity 71.4%; Pred. No. 91;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
|||:||||
Db 217 VLDVAEF 223

RESULT 11
T18315
hypothetical protein L7610.5 - Leishmania major
C;Species: Leishmania major
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T18315
R;Oliver, K.; Murphy, L.; Quail, M.; Lawson, D.; Harris, D.; Rajandream, M.; Ivens, A.;
submitted to the EMBL Data Library, May 1999
A;Reference number: Z18876
A;Accession: T18315
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-248 <OLI>
A;Cross-references: UNIPROT:O97008; EMBL:AL034356; NID:ei371878; PID:ei371560; PIDN:CAA2
C;Genetics:
A;Note: L7610.5

Query Match 82.8%; Score 24; DB 2; Length 248;
Best Local Similarity 71.4%; Pred. No. 95;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
|||:||||
Db 196 LMRVAEF 202

RESULT 12
G81693
serine/threonine protein phosphatase, probable TC0530 [imported] - Chlamydia muridarum (C
C;Species: Chlamydia muridarum, Chlamydia trachomatis MoPn
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C;Accession: G81693
R;Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey, J.
C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,
Nucleic Acids Res. 28, 1397-1406, 2000
A;Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.
A;Reference number: A81500; MUID:20150255; PMID:10684935
A;Accession: G81693
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-249 <TET>
A;Cross-references: UNIPROT:Q9PKD5; GB:AE002321; GB:AE002160; NID:g7190560; PIDN:AAF3937

A;Experimental source: strain Nigg (MoPn)
C;Genetics:
A;Gene: TC0530
C;Superfamily: conserved hypothetical protein yloO; conserved hypothetical protein yloO

Query Match 82.8%; Score 24; DB 2; Length 249;
Best Local Similarity 71.4%; Pred. No. 96;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
: |||||
: |||||

Db 1 MMIVAEF 7

RESULT 13
F95103
6-phosphofructokinase [imported] - Streptococcus pneumoniae (strain TIGR4)
C;Species: Streptococcus pneumoniae
C;Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 09-Jul-2004
C;Accession: F95103
R;Tettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid
on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapf
le, S.; Hickey, E.K.; Holt, I.E.
Science 293, 498-506, 2001
A;Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,
A;Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
A;Reference number: A95000; MUID:21357209; PMID:11463916
A;Accession: F95103
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-335 <KUR>
A;Cross-references: UNIPROT:Q97RC6; GB:AE005672; PIDN:AAK75023.1; PID:gl14972371; GSPDB:G
A;Experimental source: strain TIGR4
C;Genetics:
A;Gene: SP0896
C;Superfamily: ATP-dependent phosphofructokinase, prokaryotic type; 6-phosphofructokinase

Query Match 82.8%; Score 24; DB 2; Length 335;
Best Local Similarity 71.4%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
: |||||
: |||||

Db 224 VMSAAEF 230

RESULT 14
D97971
6-phosphofructokinase (EC 2.7.1.11) [imported] - Streptococcus pneumoniae (strain R6)
C;Species: Streptococcus pneumoniae
C;Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 09-Jul-2004
C;Accession: D97971
R;Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; B
e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M
y, P.; Sun, P.M.; Winkler, M.E.
J. Bacteriol. 183, 5709-5717, 2001
A;Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;
A;Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
A;Reference number: A97872; MUID:21429245; PMID:11544234
A;Accession: D97971
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-335 <KUR>
A;Cross-references: UNIPROT:Q8DQ85; GB:AE007317; PIDN:AAK99600.1; PID:gl15458395; GSPDB:G
A;Genetics:
A;Gene: pfkA
C;Superfamily: ATP-dependent phosphofructokinase, prokaryotic type; 6-phosphofructokinase
C;Keywords: phosphotransferase

Query Match 82.8%; Score 24; DB 2; Length 335;
Best Local Similarity 71.4%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
: |||||
: |||||

Db 224 VMSAAEF 230

RESULT 15
F97010
fusion, chorismate mutase and shikimate 5-dehydrogenase [imported] - Clostridium acetobut
C;Species: Clostridium acetobutylicum
C;Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 09-Jul-2004
C;Accession: F97010
R;Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,
.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clo
A;Reference number: A96900; MUID:21359325; PMID:21359325
A;Accession: F97010
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-367 <KUR>
A;Cross-references: UNIPROT:Q97KM0; GB:AE001437; PIDN:AAK78873.1; PID:gl15023795; GSPDB:G
A;Experimental source: Clostridium acetobutylicum ATCC824
C;Genetics:
A;Gene: CAC0897

Query Match 82.8%; Score 24; DB 2; Length 367;
Best Local Similarity 71.4%; Pred. No. 1.4e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
: |||||
: |||||

Db 29 VMEVABY 35

RESULT 16
E86416
unknown protein, 31966-27982 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C;Accession: E86416
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, I
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: E86416
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-484 <STO>
A;Cross-references: UNIPROT:Q9C7R2; GB:AE005172; MID:gl0092221; PIDN:AAI2637.1; GSPDB:G
C;Genetics:
A;Map position: 1

Query Match 82.8%; Score 24; DB 2; Length 484;
Best Local Similarity 71.4%; Pred. No. 1.9e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
: |||||
: |||||

Db 437 VLQVAEF 443

RESULT 17
B72660
probable type II DNA topoisomerase VI subunit b APE0706 - Aeropyrum pernix (strain K1)
C;Species: Aeropyrum pernix
C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004

C;Accession: B72660
R;Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takahawa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; M DNA Res. 6, 83-101, 1999
A;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyrum
A;Reference number: A72450; MUID:99310339; PMID:10382966
A;Accession: B72660
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-565 <KAW>
A;Cross-references: UNIPROT:Q9YE64; DDBJ:AP0000060; NID:G5104188; PIDN:BAA79682.1; PID:G5
A;Experimental source: strain K1
C;Genetics:
A;Gene: AP00706
C;Superfamily: Archaeoglobus fulgidus probable DNA topoisomerase VI chain B

Query Match 82.8%; Score 24; DB 2; Length 565;
Best Local Similarity 83.3%; Pred. No. 2.2e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 MXVAEF 7
|||
Db 28 MSVAEF 33

RESULT 18
E71052
hypothetical protein PH116 - Pyrococcus horikoshii
C;Species: Pyrococcus horikoshii
C;Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 09-Jul-2004
C;Accession: E71052
R;Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Sekir M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi DNA Res. 5, 55-76, 1998
A;Title: Complete sequence and gene organization of the genome of a hyper-thermophilic a
A;Reference number: A71000; MUID:98344137; PMID:9679194
A;Accession: E71052
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-582 <KAW>
A;Cross-references: UNIPROT:O58843; GB:AP0000005; NID:G3236132; PIDN:BAA30215.1; PID:d103
A;Experimental source: strain OT3
A;Note: this accession replaces an interim accession for a sequence replaced by GenBank
C;Genetics:
A;Gene: PH116

Query Match 82.8%; Score 24; DB 2; Length 582;
Best Local Similarity 57.1%; Pred. No. 2.2e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
|||
Db 440 VMAIEY 446

RESULT 19
F75090
archaeosine trna-ribosyltransferase (tgta) PAB0740 - Pyrococcus abyssi (strain Orsay)
C;Species: Pyrococcus abyssi
C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
C;Accession: F75090
R;anonymous, Genoscope
A;Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome stru
A;Reference number: A75001
A;Accession: F75090
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-584 <KAW>
A;Cross-references: UNIPROT:Q9UZN0; GB:AJ248286; GB:AL096836; NID:G5458366; PIDN:CAB5002
A;Experimental source: strain Orsay
C;Genetics:
A;Gene: tgta; PAB0740

Query Match 82.8%; Score 24; DB 2; Length 584;
Best Local Similarity 57.1%; Pred. No. 2.2e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
|||
Db 440 VMAIEY 446

RESULT 20
T25551
hypothetical protein C17H11.2 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T25551
R;Johnson, D.
A;Description: The sequence of C. elegans cosmid C17H11.
A;Reference number: Z20049
A;Accession: T25551
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-754 <JOH>
A;Cross-references: UNIPROT:P91063; EMBL:U80847; PIDN:AAB37983.1; GSPDB:GN000028; CESP:C17
A;Experimental source: strain Bristol N2; clone C17H11
C;Genetics:
A;Gene: CESP.C17H11.2
A;Map position: X
A;Introns: 155/1; 246/3; 288/3; 368/3; 631/1
C;Superfamily: Caenorhabditis elegans hypothetical protein C17H11.2

Query Match 82.8%; Score 24; DB 2; Length 754;
Best Local Similarity 83.3%; Pred. No. 2.9e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 MXVAEF 7
|||
Db 556 MTVAEF 561

RESULT 21
T29581
hypothetical protein M03A1.1 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 31-Jan-2000
C;Accession: T29581
R;Fulton, B.; Marra, M.
A;Description: The sequence of C. elegans cosmid M03A1.
A;Reference number: Z20645
A;Accession: T29581
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-919 <FUL>
A;Cross-references: EMBL:U49956; PIDN:AAA93435.1; CESP:M03A1.1
A;Gene: CESP.M03A1.1
A;Introns: 8/3; 92/1; 142/1; 404/1; 437/1; 645/3
C;Superfamily: protein-tyrosine kinase, receptor type eph; fibronectin type III repeat h

Query Match 82.8%; Score 24; DB 2; Length 919;
Best Local Similarity 71.4%; Pred. No. 3.5e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
|||
Db 592 VMVAEF 598

RESULT 22
H69643
isoleucine-tRNA ligase (EC 6.1.1.5) ileS - Bacillus subtilis

N;Alternate names: isoleucyl-tRNA synthetase

C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
 C;Accession: H69643
 R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Allioni, G.; Azevedo, V.; Bertoni
 C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chd
 A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
 Nature 390, 249-256, 1997
 A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallen
 lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.P.
 Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
 A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mauvel
 Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
 Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon,
 A;Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Segiguchi, J.; Sekowska, A.; Seron
 akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
 T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
 A;Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
 A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
 A;Reference number: A69580; MUID:98044033; PMID:9384377
 A;Accession: H69643
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-921 <KUN>
 A;Cross-references: UNIPROT:Q45477; GB:Z99112; GB:AL009126; NID:g2633902; PIDN:CAB13417.
 A;Experimental source: strain 168
 C;Genetics:
 C;Superfamily: isoleucine-tRNA ligase
 C;Keywords: aminoacyl-tRNA synthetase; ligase; protein biosynthesis

Query Match 82.8%; Score 24; DB 2; Length 921;

Best Local Similarity 83.3%; Pred. No. 3.5e+02;

Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 MXVAEF 7
 |||||
 Db 116 MSVAEP 121

RESULT 23

A83968

isoleucyl-tRNA synthetase ileS [imported] - Bacillus halodurans (strain C-125)

C;Species: Bacillus halodurans

C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004

C;Accession: A83968

R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, N.; Masui, N.; Fujii, F.; Hira

Nucleic Acids Res. 28, 4317-4331, 2000

A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and

A;Reference number: A83650; MUID:20512582; PMID:11058132

A;Accession: A83968

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-921 <STO>

A;Cross-references: UNIPROT:Q9K9V0; GB:AP001515; GB:BA000004; NID:gl0174886; PIDN:BA062

A;Experimental source: strain C-125

C;Genetics:

A;Gene: ileS

C;Superfamily: isoleucine-tRNA ligase

Query Match 82.8%; Score 24; DB 2; Length 921;

Best Local Similarity 83.3%; Pred. No. 3.5e+02;

Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 MXVAEF 7
 |||||
 Db 115 MSVAEP 120

RESULT 24

T42400

Eph receptor tyrosine kinase - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004

C;Accession: T42400

R;George, S.E.; Simokat, K.; Hardin, J.; Chisholm, A.D.

Cell 92, 633-643, 1998

A;Title: The VAB-1 Eph receptor tyrosine kinase functions in neural and epithelial morph

A;Reference number: Z22158; MUID:98165343; PMID:9506518

A;Accession: T42400

A;Status: preliminary; translated from GB/EMBL/DBBJ

A;Molecule type: DNA

A;Residues: 1-1122 <GEO>

A;Cross-references: UNIPROT:O61460; EMBL:AF040269; NID:g2828679; PIDN:AAC38970.1; PID:g28

C;Genetics:

A;Gene: vab-1

C;Superfamily: protein-tyrosine kinase, receptor type eph; fibronectin type III repeat h

Query Match 82.8%; Score 24; DB 2; Length 1122;

Best Local Similarity 71.4%; Pred. No. 4.3e+02;

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
 |||||
 Db 790 VMVAEY 796

RESULT 25

T15018

hypothetical protein Y1103 - Yersinia pestis plasmid pMT1

C;Species: Yersinia pestis

C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004

C;Accession: T15018

R;Lindler, L.E.; Plano, G.V.; Burland, V.; Mayhew, G.F.; Blattner, F.R.

Infect. Immun. 66, 5731-5742, 1998

A;Title: Complete DNA sequence and detailed analysis of the Yersinia pestis KIM5 plasmid

A;Reference number: Z18268; MUID:99043898; PMID:9826348

A;Accession: T15018

A;Status: preliminary; translated from GB/EMBL/DBBJ

A;Molecule type: DNA

A;Residues: 1-89 <LIN>

A;Cross-references: UNIPROT:Q9ZGY2; EMBL:AF074611; NID:g3883003; PID:g3883101; PIDN:AAC8

C;Genetics:

A;Gene: Y1103

A;Genome: plasmid pMT1

Query Match 79.3%; Score 23; DB 2; Length 89;

Best Local Similarity 83.3%; Pred. No. 64;

Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 MXVAEF 7
 |||||
 Db 1 MFVAEF 6

RESULT 26

AB3053

hypothetical protein Atu4039 [imported] - Agrobacterium tumefaciens (strain C58, Dupont)

C;Species: Agrobacterium tumefaciens

C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004

C;Accession: AB3053

R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L.

erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClellan

; Karp, P.; Romero, P.; Zhang, S.

Science 294, 2317-2323, 2001

A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, M.

ster, E.W.

A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.

A;Reference number: AB2577; MUID:21608550; PMID:11743193

A;Accession: AB3053

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-96 <KUR>

A;Cross-references: UNIPROT:Q8U8Q1; GB:AE008689; PIDN:AAL44840.1; PID:gl17742484; GSPDB:G

A;Experimental source: strain C58 (Dupont)

C;Genetics:

A:Gene: Atu4039

A:Map position: linear chromosome

Query Match 79.3%; Score 23; DB 2; Length 96;

Best Local Similarity 57.1%; Pred. No. 69;

Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7

:: ||||

Db 52 ILNVAEF 58

RESULT 27

E90987

hypothetical protein ECs2869 [imported] - Escherichia coli (strain O157:H7, substrain R)

C:Species: Escherichia coli

C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004

C:Accession: E90987

R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.C.

gaawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.

DNA Res. 8, 11-22, 2001

A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gen

A:Reference number: A99629; MUID:21156231; PMID:11258796

A:Accession: E90987

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-112 <HAY>

A:Cross-references: UNIPROT:Q8X7L7; GB:BA000007; PIDN:BA36292.1; PID:gl3362338; GSPDB:C

A:Experimental source: strain O157:H7, substrain RIMD 0509952

C:Genetics:

A:Gene: ECs2869

Query Match 79.3%; Score 23; DB 2; Length 112;

Best Local Similarity 71.4%; Pred. No. 80;

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7

:: ||||

Db 66 VMTVAKF 72

RESULT 28

H85832

hypothetical protein Z3230 [imported] - Escherichia coli (strain O157:H7, substrain EDL9

C:Species: Escherichia coli

C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004

C:Accession: H85832

R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew

iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,

Nature 409, 529-533, 2001

A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.

A:Reference number: A85480; MUID:21074935; PMID:11206551

A:Accession: H85832

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-112 <STO>

A:Cross-references: UNIPROT:Q8X7L7; GB:AE005174; NID:gl216261; PIDN:AAG57124.1; GSPDB:C

A:Experimental source: strain O157:H7, substrain EDL933

C:Genetics:

A:Gene: Z3230

Query Match 79.3%; Score 23; DB 2; Length 112;

Best Local Similarity 71.4%; Pred. No. 80;

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7

:: ||||

Db 66 VMTVAKF 72

RESULT 29

H69515

riboflavin synthase, subunit beta (ribe) homolog - Archaeoglobus fulgidus

C:Species: Archaeoglobus fulgidus

C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004

C:Accession: H69515

R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson,

; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.;

Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.

Nature 390, 364-370, 1997

A:Authors: Uterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.N.

Smith, H.O.; Woese, C.R.; Venter, J.C.

A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo

A:Reference number: A69250; MUID:98049343; PMID:9389475

A:Accession: H69515

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-143 <KLE>

A:Cross-references: UNIPROT:Q28152; GB:AE000957; GB:AE000782; NID:g2689280; PIDN:AA889124

C:Superfamily: riboflavin synthase beta chain

Query Match 79.3%; Score 23; DB 2; Length 143;

Best Local Similarity 83.3%; Pred. No. 1e+02;

Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 MVVAEF 7

:: ||||

Db 8 MVVAEF 13

RESULT 30

G00016

FGF-receptor - common marmoset (fragment)

C:Species: Callithrix jacchus (common marmoset)

C:Date: 13-Mar-1997 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004

C:Accession: G00016

R:Einspanner, R.

submitted to the EMBL Data Library, December 1995

A:Reference number: H00018

A:Accession: G00016

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-157 <EIN>

A:Cross-references: UNIPROT:Q28332; EMBL:Z68149; NID:gl279349

C:Superfamily: basic fibroblast growth factor receptor 1; immunoglobulin homology; prote

F;1-157/Domain: protein kinase homology (fragment) <KIN>

Query Match 79.3%; Score 23; DB 2; Length 157;

Best Local Similarity 57.1%; Pred. No. 1.1e+02;

Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7

:: ||||

Db 106 VMKIADF 112

RESULT 31

G86842

shikimate kinase (EC 2.7.1.71) [imported] - Lactococcus lactis subsp. lactis (strain IL14

C:Species: Lactococcus lactis subsp. lactis

C:Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 09-Jul-2004

C:Accession: G86842

R:Boletín, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarre, K.; Weissenbach, J.; Ehrlic

Genome Res. 11, 731-753, 2001

A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis ssp

A:Reference number: A86625; MUID:21235186; PMID:11337471

A:Accession: G86842

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-162 <STO>

A:Cross-references: UNIPROT:Q9CEU1; GB:AE005176; PID:gl2724763; PIDN:AAK05841.1; GSPDB:G

A:Experimental source: strain IL1403

C:Genetics:

A:Gene: aroK

C:Superfamily: shikimate kinase; shikimate kinase homology

C:Keywords: phosphotransferase

```

Query Match          79.3%; Score 23; DB 2; Length 162;
Best Local Similarity 66.1%; Pred. No. 1.2e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 MVXAEF 7
   |||
Db 37 MTIAEF 42

RESULT 32
AB2739
acetyltransferase [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C;Species: Agrobacterium tumefaciens
C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
C;Accession: AB2739
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L.;
Ergae, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClellan,
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A;Reference number: AB2577; MUID:21608550; PMID:11743193
A;Accession: AB2739
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-169 <KUR>
A;Cross-references: UNIPROT:Q8UFS5; GB:AE008688; PIDN:AAL42328.1; PID:g17739732; GSPDB:G
A;Experimental source: strain C58 (Dupont)
C;Genetics:
A;Gene: Atul322
A;Map position: circular chromosome

Query Match          79.3%; Score 23; DB 2; Length 169;
Best Local Similarity 83.3%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 MVXAEF 7
   |||
Db 55 MRVAEF 60

RESULT 33
S27019
fibroblast growth factor receptor - Japanese medaka (fragment)
N;Contains: protein-tyrosine kinase (EC 2.7.1.112)
C;Species: Oryzias latipes (Japanese medaka)
C;Date: 19-Mar-1997 #sequence_revision 24-Jul-1997 #text_change 09-Jul-2004
C;Accession: S27019
R;Emori, Y.; Yasuoka, A.; Saigo, K.
FEBS Lett. 314, 176-178, 1992
A;Title: Identification of four FGF receptor genes in Medaka fish (Oryzias latipes).
A;Reference number: S27019; MUID:93093167; PMID:1459248
A;Accession: S27019
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-172 <EMO>
A;Cross-references: UNIPROT:Q91176; GB:D13550; NID:g222932; PIDN:BAA02749.1; PID:g222933
C;Superfamily: basic fibroblast growth factor receptor 1; immunoglobulin homology; protei
C;Keywords: alternative splicing; ATP; autophosphorylation; growth factor receptor; magn
F;1-172/Domain: protein kinase homology (fragment) <KIN>
F;23,40,132/Active site: Lys, Glu, Asp #status predicted
F;137,150/Binding site: magnesium (Asn, Asp) #status predicted
F;163/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status predicted

Query Match          79.3%; Score 23; DB 2; Length 172;
Best Local Similarity 57.1%; Pred. No. 1.2e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
   |||
Db 145 VMKIADF 151

RESULT 36
S27020
fibroblast growth factor receptor - Japanese medaka (fragment)
N;Contains: protein-tyrosine kinase (EC 2.7.1.112)
C;Species: Oryzias latipes (Japanese medaka)
C;Date: 19-Mar-1997 #sequence_revision 24-Jul-1997 #text_change 09-Jul-2004
C;Accession: S27020

```

```

RESULT 34
S27022
fibroblast growth factor receptor - Japanese medaka (fragment)
N;Contains: protein-tyrosine kinase (EC 2.7.1.112)
C;Species: Oryzias latipes (Japanese medaka)
C;Date: 19-Mar-1997 #sequence_revision 24-Jul-1997 #text_change 09-Jul-2004
C;Accession: S27022
R;Emori, Y.; Yasuoka, A.; Saigo, K.
FEBS Lett. 314, 176-178, 1992
A;Title: Identification of four FGF receptor genes in Medaka fish (Oryzias latipes).
A;Reference number: S27019; MUID:93093167; PMID:1459248
A;Accession: S27022
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-172 <EMO>
A;Cross-references: UNIPROT:Q02529; GB:D13553; NID:g222938; PIDN:BAA02752.1; PID:g222939
C;Superfamily: basic fibroblast growth factor receptor 1; immunoglobulin homology; protei
C;Keywords: alternative splicing; ATP; autophosphorylation; growth factor receptor; magn
F;1-172/Domain: protein kinase homology (fragment) <KIN>
F;23,40,132/Active site: Lys, Glu, Asp #status predicted
F;137,150/Binding site: magnesium (Asn, Asp) #status predicted
F;163/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status predicted

Query Match          79.3%; Score 23; DB 2; Length 172;
Best Local Similarity 57.1%; Pred. No. 1.2e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
   |||
Db 145 VMKIADF 151

RESULT 35
S27021
fibroblast growth factor receptor - Japanese medaka (fragment)
N;Contains: protein-tyrosine kinase (EC 2.7.1.112)
C;Species: Oryzias latipes (Japanese medaka)
C;Date: 19-Mar-1997 #sequence_revision 29-Aug-1997 #text_change 09-Jul-2004
C;Accession: S27021
R;Emori, Y.; Yasuoka, A.; Saigo, K.
FEBS Lett. 314, 176-178, 1992
A;Title: Identification of four FGF receptor genes in Medaka fish (Oryzias latipes).
A;Reference number: S27019; MUID:93093167; PMID:1459248
A;Accession: S27021
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-172 <EMO>
A;Cross-references: UNIPROT:Q02528; GB:D13552; NID:g222936; PIDN:BAA02751.1; PID:g222937
C;Superfamily: basic fibroblast growth factor receptor 1; immunoglobulin homology; protei
C;Keywords: alternative splicing; ATP; autophosphorylation; growth factor receptor; magn
F;1-172/Domain: protein kinase homology (fragment) <KIN>
F;23,40,132/Active site: Lys, Glu, Asp #status predicted
F;137,150/Binding site: magnesium (Asn, Asp) #status predicted
F;163/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status predicted

Query Match          79.3%; Score 23; DB 2; Length 172;
Best Local Similarity 57.1%; Pred. No. 1.2e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
   |||
Db 145 VMKIADF 151

RESULT 36
S27020
fibroblast growth factor receptor - Japanese medaka (fragment)
N;Contains: protein-tyrosine kinase (EC 2.7.1.112)
C;Species: Oryzias latipes (Japanese medaka)
C;Date: 19-Mar-1997 #sequence_revision 24-Jul-1997 #text_change 09-Jul-2004
C;Accession: S27020

```


R;Emori, Y.; Yasuoka, A.; Saigo, K.
FEBS Lett. 314, 176-178, 1992
A;Title: Identification of four FGF receptor genes in Medaka fish (*Oryzias latipes*).
A;Reference number: S27019; MUID:93093167; PMID:1459248
A;Accession: S27020
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-172 <EMO>
A;Cross-references: UNIPROT:Q91177; GB:D13551; NID:G222934; PIDN:BAA02750.1; PID:G222934
C;Superfamily: basic fibroblast growth factor receptor 1; immunoglobulin homology; protein
C;Keywords: alternative splicing; ATP; autophosphorylation; growth factor receptor; magn
F;1-172/Domain: protein kinase homology (fragment) <KIN>
F;23,40,132/Active site: Lys, Glu, Asp #status predicted
F;137,150/Binding site: magnesium (Asn, Asp) #status predicted
F;163/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status predicted

Query Match 79.3%; Score 23; DB 2; Length 172;
Best Local Similarity 57.1%; Pred. No. 1.2e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
Db 145 VMKIADF 151

RESULT 37
S16314
Photosynthetic reaction center cytochrome c chain precursor - *Erythrobacter* sp. (fragment)
C;Species: *Erythrobacter* sp.
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 09-Jul-2004
C;Accession: S16314
R;Liebetanz, R.; Hornberger, U.; Drews, G.
Mol. Microbiol. 5, 1459-1468, 1991
A;Title: Organization of the genes coding for the reaction-centre L and M subunits and E

Query Match 79.3%; Score 23; DB 2; Length 187;
Best Local Similarity 83.3%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 MXVAEF 7
Db 58 MRVAEF 63

RESULT 38
A82392
transcription regulator TetR family VCA0999 [imported] - *Vibrio cholerae* (strain N16961)
C;Species: *Vibrio cholerae*
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C;Accession: A82392
R;Heidelberg, J.P.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Basu, S.; Qin, H.; Dragoi, I.; Sellers, F.
L. R.K.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A;Title: DNA Sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.
A;Reference number: A82035; MUID:20406833; PMID:10952301
A;Accession: A82392
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-191 <HEI>
A;Cross-references: UNIPROT:Q9KKV1; GB:AE004426; GB:AE003853; NID:G9658431; PIDN:AAF9689

A;Experimental source: serogroup O1; strain N16961; biotype El Tor
C;Genetics:
A;Gene: VCA0999
A;Map position: 2

Query Match 79.3%; Score 23; DB 2; Length 191;
Best Local Similarity 83.3%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 MXVAEF 7
Db 1 MRVAEF 6

RESULT 39
H97519
hypothetical protein AGR_C_2436 [imported] - *Agrobacterium tumefaciens* (strain C58, Cere
C;Species: *Agrobacterium tumefaciens*
C;Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004
C;Accession: H97519
R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,
A.; Liu, F.; Wollam, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent *Agrobacterium tum*
A;Reference number: A97359; MUID:21608551; PMID:11743194
A;Accession: H97519
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-209 <KUR>
A;Cross-references: UNIPROT:Q8UF85; GB:AE007869; PIDN:AAK87113.1; PID:GL5156377; GSPDB:G
C;Genetics:
A;Gene: AGR_C_2436
A;Map position: circular chromosome

Query Match 79.3%; Score 23; DB 2; Length 209;
Best Local Similarity 83.3%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 MXVAEF 7
Db 95 MRVAEF 100

RESULT 40
D81436
probable tonB transport protein Cj0181 [imported] - *Campylobacter jejuni* (strain NCTC 11
C;Species: *Campylobacter jejuni*
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C;Accession: D81436
R;Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chilling
C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVliet, A.; Whitehead, S.; Barre
Nature 403, 665-668, 2000
A;Title: The genome sequence of the food-borne pathogen *Campylobacter jejuni* reveals hyp
A;Reference number: A81250; MUID:20150912; PMID:10688204
A;Accession: D81436
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-248 <FAR>
A;Cross-references: UNIPROT:Q9PIU8; GB:AL139074; GB:AL111168; NID:G9697505; PIDN:CAB7266
A;Experimental source: serotype O2, strain NCTC 11168
C;Genetics:
A;Gene: tonB1; Cj0181

Query Match 79.3%; Score 23; DB 2; Length 248;
Best Local Similarity 57.1%; Pred. No. 1.8e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
Db 47 IMMVSEF 53

RESULT 41

BVBV53
phosphomannomutase (EC 5.4.2.8) - yeast (Saccharomyces cerevisiae)
N;Alternate names: protein YFL045c
C;Species: Saccharomyces cerevisiae
C;Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 09-Jul-2004
C;Accession: S05874; S56210
R;Bernstein, M.; Hoffmann, W.; Ammerer, G.; Schekman, R.
J. Cell Biol. 101, 2374-2382, 1985
A;Title: Characterization of a gene product (sec53p) required for protein assembly in the yeast Saccharomyces cerevisiae
A;Reference number: S05874; MUID:86059690; PMID:3905826
A;Accession: S05874
A;Molecule type: DNA
A;Residues: 1-254 <BER>
A;Cross-references: UNIPROT:P07283; EMBL:X03213; NID:g4444; PIDN:CAA26957.1; PID:g4445
R;Murakami, Y.; Naitou, M.; Hagiwara, H.; Shibata, T.; Ozawa, M.; Sasanuma, S.I.; Sasano, Y. Submitted to the EMBL Data Library, May 1995
A;Description: Analysis of the nucleotide sequence of chromosome VI from Saccharomyces cerevisiae
A;Reference number: S56186
A;Accession: S56210
A;Molecule type: DNA
A;Residues: 1-254 <MUR>
A;Cross-references: EMBL:D50617; NID:g836685; PIDN:BAA09196.1; PID:g836710; GSPDB:GN000000000
C;Genetics:
A;Gene: SGD:SEC53; MIPS:YFL045C
A;Cross-references: SGD:S0001849; MIPS:YFL045C
A;Map position: 6L
C;Superfamily: SEC53 protein
C;Keywords: intramolecular transferase; isomerase

Query Match 79.3%; Score 23; DB 1; Length 254;
Best Local Similarity 66.7%; Pred. No. 1.8e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 MVVAEF 7
| : |||
Db 1 MSIAEF 6

RESULT 42
S28749
NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 1 - blue mussel mitochondrion (fragment)
C;Species: mitochondrion Mytilus edulis (blue mussel)
C;Date: 22-Nov-1993 #sequence_revision 26-Jul-1996 #text_change 09-Jul-2004
C;Accession: S28749; S28750
R;Hoffmann, R.J.; Boore, J.D.; Brown, W.M.
Genetics 131, 397-412, 1992
A;Title: A novel mitochondrial genome organization for the blue mussel, Mytilus edulis.
A;Reference number: S28743; MUID:92354892; PMID:1386586
A;Accession: S28749
A;Molecule type: DNA
A;Residues: 1-104;105-275 <BRO>
A;Cross-references: UNIPROT:Q00860; EMBL:M83758; EMBL:M83759
C;Genetics:
A;Gene: ND1
A;Genome: mitochondrion
A;Genetic code: SGC4
A;Start codon: GTG
C;Superfamily: NADH:quinone oxidoreductase (complex I), subunit 1/[NiFe]-hydrogenase-3-b
C;Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation;

Query Match 79.3%; Score 23; DB 2; Length 275;
Best Local Similarity 57.1%; Pred. No. 2e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
| : |||
Db 196 VMFTAEY 202

RESULT 43
G69479
methionyl aminopeptidase (map) homolog - Archaeoglobus fulgidus
C;Species: Archaeoglobus fulgidus

C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C;Accession: G69479
R;Kienk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson, R.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.; Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L. Nature 390, 364-370, 1997
A;Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.A. Smith, H.O.; Woese, C.R.; Venter, J.C.
A;Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeon Pyrococcus furiosus
A;Reference number: A69250; MUID:98049343; PMID:9389475
A;Accession: G69479
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-291 <KLE>
A;Cross-references: UNIPROT:Q28438; GB:AE000976; GB:AE000782; NID:g2689299; PIDN:AB889413
C;Superfamily: Escherichia coli methionyl aminopeptidase

Query Match 79.3%; Score 23; DB 2; Length 291;
Best Local Similarity 57.1%; Pred. No. 2.1e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
| : |||
Db 33 ILEVAEF 39

RESULT 44
F36819
C14 protein - rabbit fibroma virus
C;Species: rabbit fibroma virus, Shope fibroma virus
C;Date: 23-Apr-1993 #sequence_revision 23-Apr-1993 #text_change 17-Nov-2000
C;Accession: F36819
R;Strayer, D.S.; Jerng, H.H.; O'Connor, K.
Virology 185, 585-595, 1991
A;Title: Sequence and analysis of a portion of the genomes of Shope fibroma virus and mouse mammary tumor virus
A;Reference number: A41700; MUID:92074222; PMID:1660196
A;Accession: F36819
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-370 <STR>
A;Cross-references: GB:M32743
C;Superfamily: vaccinia virus nucleoside-triphosphatase I

Query Match 79.3%; Score 23; DB 2; Length 370;
Best Local Similarity 71.4%; Pred. No. 2.6e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
| : |||
Db 158 VQSVAEF 164

RESULT 45
AC0253
ribonuclease III (EC 3.1.26.3) [imported] - Yersinia pestis (strain CO92)
C;Species: Yersinia pestis
C;Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004
C;Accession: AC0253
R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.; deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.; Eil, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrall, E. Nature 413, 523-527, 2001
A;Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A;Reference number: AB0001; MUID:21470413; PMID:11586360
A;Accession: AC0253
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-373 <KUR>
A;Cross-references: UNIPROT:Q8ZES8; GB:AL590842; PIDN:CAC90887.1; PID:g15980086; GSPDB:G69479
C;Genetics:
A;Gene: rnd
C;Superfamily: ribonuclease D

Query Match 79.3%; Score 23; DB 2; Length 373;
Best Local Similarity 83.3%; Pred. No. 2.6e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 MXVAEF 7
| | | | |
Db 120 MLVAEF 125

RESULT 46

A48352 Genome polypeptide - tick-borne encephalitis virus (strain K23) (fragment)

N;Contains: membrane protein M; nonstructural protein NS1
C;Species: tick-borne encephalitis virus
C;Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 09-Jul-2004
C;Accession: A48352; S18104
R;Jacobs, S.C.; Stephenson, J.R.; Wilkinson, G.W.
Arch. Virol. 129, 279-285, 1993
A;Title: Sequence of the NS 1 gene of the K 23 isolate of tick-borne encephalitis virus
A;Reference number: A48352; MUID:93228450; PMID:8470955
A;Accession: A48352
A;Molecule type: genomic RNA
A;Residues: 1-377 <JAC>
A;Cross-references: UNIPROT:Q88482; EMBL:X62886; NID:g62032; PIDN:CAA44678.1; PID:g62033
A;Note: sequence extracted from NCBI backbone (NCBIP:129078)
C;Superfamily: yellow fever virus genome polypeptide
C;Keywords: membrane protein; nonstructural protein; polypeptide

Query Match 79.3%; Score 23; DB 2; Length 377;
Best Local Similarity 71.4%; Pred. No. 2.7e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
| | | | |
Db 178 VFTVAEF 184

RESULT 47

H83223 conserved hypothetical protein PA3374 [imported] - Pseudomonas aeruginosa (strain PA01)

C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warriner, P.; Hickey, M.J.; B
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
-; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A;Reference number: A82950; MUID:20437337; PMID:10984043
A;Accession: H83223
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-387 <STO>
A;Cross-references: UNIPROT:Q9HYM7; GB:AE004759; GB:AE004091; NID:g9949500; PIDN:AAG0676
A;Experimental source: strain PA01
C;Genetics:
A;Gene: PA3374

Query Match 79.3%; Score 23; DB 2; Length 387;
Best Local Similarity 66.7%; Pred. No. 2.7e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 MXVAEF 7
| | | | |
Db 249 MAIAEF 254

RESULT 48

I51023 fibroblast growth factor receptor 1 - eastern newt (fragment)

C;Species: Notophthalmus viridescens, Triturus viridescens (eastern newt)
C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 09-Jul-2004
C;Accession: I51023

R;Poulin, M.L.; Patrie, K.M.; Botelho, M.J.; Tassava, R.A.; Chiu, I.
Development 119, 353-361, 1993

A;Title: Heterogeneity in the expression of fibroblast growth factor receptors during lin
A;Reference number: I51023; MUID:94116439; PMID:8287792

A;Accession: I51023
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-388 <POU>
A;Cross-references: UNIPROT:Q91146; GB:L19868; NID:g476726; PIDN:AAA49394.1; PID:g476727
C;Genetics:

A;Gene: FGFR1
C;Superfamily: basic fibroblast growth factor receptor 1; immunoglobulin homology; protei
C;Keywords: ATP; growth factor receptor
F;44-329/Domain: protein kinase homology <KIN>
F;52-60/Region: protein kinase ATP-binding motif

Query Match 79.3%; Score 23; DB 2; Length 388;
Best Local Similarity 57.1%; Pred. No. 2.8e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
| | | | |
Db 204 VMKIADF 210

RESULT 49

H75137 probable carbon-sulfur lyase (EC 4.4.1.-) PAB0605 - Pyrococcus abyssi (strain Orsay)
C;Species: Pyrococcus abyssi
C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
C;Accession: H75137
R;anonymous, Genoscope
submitted to the EMBL Data Library, July 1999

A;Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome stru
A;Reference number: A75001
A;Accession: H75137
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-393 <KAW>
A;Cross-references: UNIPROT:Q9V088; GB:AJ248285; GB:AL096836; NID:g5458067; PIDN:CAB4981;
A;Experimental source: strain Orsay
C;Genetics:
A;Gene: metB; PAB0605
C;Superfamily: O-succinylhomoserine (thiol)-lyase
C;Keywords: carbon-oxygen lyase; carbon-sulfur lyase

Query Match 79.3%; Score 23; DB 2; Length 393;
Best Local Similarity 66.7%; Pred. No. 2.8e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 MXVAEF 7
| | | | |
Db 268 MAIAEF 273

RESULT 50

A11302 pantothenate metabolism flavoprotein homolog lmo1825 [imported] - Listeria monoc
C;Species: Listeria monocytogenes
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C;Accession: A11302

R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker,
-; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Feihi, H.,
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001

A;Authors: Kreft, J.; Kuhn, M.; Kunat, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Mat
ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A;Title: Comparative Genomics of Listeria species.

A;Reference number: AB1077; MUID:21537279; PMID:11679669
A;Accession: A11302
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-399 <GLA>

A;Cross-references: UNIPROT:Q8Y674; GB:NC_003210; PIDN:CAC99903.1; PID:g16411279; GSPDB:
A;Experimental source: strain EGD-e
C;Genetics:
A;Gene: lnc1825
C;Superfamily: pantothenate metabolism flavoprotein dfp

Query Match 79.3%; Score 23; DB 2; Length 399;
Best Local Similarity 57.1%; Pred. No. 2.8e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
| : |||
Db 171 VLRIAEF 177

RESULT 51
A11674
pantothenate metabolism flavoprotein homolog lin1939 [imported] - Listeria innocua
C;Species: Listeria innocua
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C;Accession: A11674
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlant,
A;Title: Comparative genomics of Listeria species.
A;Reference number: AB1077; MUID:21537279; PMID:11679669
A;Accession: A11674
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-399 <GLA>
A;Cross-references: UNIPROT:Q92A13; GB:AL592022; PIDN:CAC97169.1; PID:g16414440; GSPDB:G
A;Experimental source: strain Clp11262
C;Genetics:
A;Gene: lin1939
C;Superfamily: pantothenate metabolism flavoprotein dfp

Query Match 79.3%; Score 23; DB 2; Length 399;
Best Local Similarity 57.1%; Pred. No. 2.8e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
| : |||
Db 171 VLRIAEF 177

RESULT 52
I65223
heparin-binding fibroblast growth factor receptor 2 - rat (fragment)
C;Species: Rattus norvegicus (Norway rat)
C;Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 09-Jul-2004
C;Accession: I65223; I65226
R;Van, G.; McBride, G.; McKeenan, W.L.
Biochem. Biophys. Res. Commun. 194, 512-518, 1993
A;Title: Exon skipping causes alteration of the COOH-terminus and deletion of the phosph
A;Reference number: I52281; MUID:93326167; PMID:8333865
A;Accession: I65223
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-415 <RES>
A;Cross-references: UNIPROT:Q64334; GB:L19106; NID:g310144; PIDN:AAB02867.1; PID:g310145
A;Accession: I65226
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-415 <BE2>
A;Cross-references: GB:L19109; NID:g310144; PIDN:AAB02867.1; PID:g310145
C;Superfamily: Basic fibroblast growth factor receptor 1; immunoglobulin homology; prote
C;Keywords: ATP; growth factor receptor; heparin binding
F;73-358/Domain: protein kinase homology <KIN>
F;81-89/Region: protein kinase ATP-binding motif

Query Match 79.3%; Score 23; DB 2; Length 415;
Best Local Similarity 57.1%; Pred. No. 2.9e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
| : |||
Db 233 VMKIADF 239

RESULT 53
A90003
hypothetical protein mura [imported] - Staphylococcus aureus (strain N315)
C;Species: Staphylococcus aureus
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C;Accession: A90003
R;Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.; I
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A;Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A;Reference number: A89758; MUID:21311952; PMID:11418146
A;Accession: A90003
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-421 <KUR>
A;Cross-references: UNIPROT:Q99SP8; GB:BA000018; PID:g13701894; PIDN:BA043186.1; GSPDB:G
A;Experimental source: strain N315
C;Genetics:
A;Gene: mura
C;Superfamily: UDP-N-acetylglucosamine 1-carboxyvinyltransferase MurZ

Query Match 79.3%; Score 23; DB 2; Length 421;
Best Local Similarity 83.3%; Pred. No. 3e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 MXVAEF 7
| : |||
Db 336 MHVAEF 341

RESULT 54
T28015
hypothetical protein ZK822.4 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T28015
R;Harris, B.
submitted to the EMBL Data Library, May 1996
A;Reference number: Z20456
A;Accession: T28015
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-445 <WIL>
A;Cross-references: UNIPROT:Q23615; EMBL:Z73898; PIDN:CAA98068.1; GSPDB:GN000022; CESP:ZK
A;Experimental source: clone ZK822
C;Genetics:
A;Gene: CESP:ZK822.4
A;Map position: 4
A;Introns: 20/2; 56/2; 69/2; 87/2; 105/2; 123/2; 254/3; 272/3; 350/3

Query Match 79.3%; Score 23; DB 2; Length 445;
Best Local Similarity 66.7%; Pred. No. 3.2e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 MXVAEF 7
| : |||
Db 169 MTIAEF 174

RESULT 55
B97213
fAD/FMN-containing dehydrogenase [imported] - Clostridium acetobutylicum
C;Species: Clostridium acetobutylicum

A;Gene: BMEI1382
A;Map position: I
C;Superfamily: NAD-dependent aldehyde dehydrogenase; aldehyde dehydrogenase homology
C;Keywords: oxidoreductase

Query Match 79.3%; Score 23; DB 2; Length 500;
Best Local Similarity 71.4%; Pred. NO. 3.5e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
|||
DB 402 VMSVLEF 408
|||

RESULT 58
AI0673
probable hydrolase STY1503 [imported] - Salmonella enterica subsp. enterica serovar Typhi
C;Species: Salmonella enterica subsp. enterica serovar Typhi
A;Title: this species has also been called Salmonella typhi
C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 16-Aug-2004
C;Accession: AI0673
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
S.; Moutle, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serovar
A;Reference number: AB0502; MUID:21534947; PMID:11677608
A;Accession: AI0673
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-594 <PAR>
A;Cross-references: GB:AL513382; PIDN:CAD01762.1; PID:gi6502610; GSPDB:GN00176
C;Genetics:
A;Gene: STY1503
C;Superfamily: Glycosyltrehalose trehalohydrolase

Query Match 79.3%; Score 23; DB 2; Length 594;
Best Local Similarity 71.4%; Pred. NO. 4.2e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
|||
DB 149 VMPVAQF 155
|||

RESULT 59
T33240
hypothetical protein T10H9.5 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 16-Aug-2004
C;Accession: T33240
R;Greco, T.; Bradshaw, H.; O'Brien, D.
submitted to the EMBL Data Library, May 1998
A;Description: The sequence of C. elegans cosmid T10H9.
A;Reference number: Z21306
A;Accession: T33240
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-598 <GRE>
A;Cross-references: UNIPROT:O76414; EMBL:AF067949; PIDN:AAC19238.1; GSPDB:GN000023; CESP:
A;Experimental source: strain Bristol N2; clone T10H9
C;Genetics:
A;Gene: CESP:T10H9.5
A;Map position: 5
A;Introns: 49/1; 91/3; 139/2; 179/2; 221/2; 270/1; 311/3; 342/2; 399/1; 481/1; 510/2; 555/2
C;Superfamily: ATP-binding cassette homology

Query Match 79.3%; Score 23; DB 2; Length 598;
Best Local Similarity 57.1%; Pred. NO. 4.2e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7

Db	340	IMRVSEF 346	:	
RESULT 60				
T49577				
hypothetical protein B208.250 [imported] - Neurospora crassa				
C:Species: Neurospora crassa				
C>Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 09-Jul-2004				
R;Schulte, U.; Algn, V.; Hoheisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,				
submitted to the Protein Sequence Database, May 2000				
A:Reference number: Z25022				
A:Accession: T49577				
A>Status: preliminary				
A:Molecule type: DNA				
A:Residues: 1-604 <SCH>				
A:Cross-references: UNIPROT:Q9P604; EMBL:AL355930; GSPDB:GN00116; NCSP:B208.250				
A:Experimental source: BAC clone B208; strain OR74A				
C:Genetics:				
A:Gene: NCSP:B208.250				
A:Map position: 6				
A:Introns: 169/2				
C:Superfamily: Neurospora crassa hypothetical protein B208.250				
Query Match				
Best Local Similarity 79.3%; Score 23; DB 2; Length 604;				
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;				
QY	1	VMXVAEF 7		
Db	384	VMPLAEF 390		
RESULT 61				
B75257				
arginyl-tRNA synthetase - Deinococcus radiodurans (strain R1)				
C:Species: Deinococcus radiodurans				
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004				
A:Accession: B75257				
R;White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;				
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma				
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.				
Science 286, 1571-1577, 1999				
A>Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.				
A:Reference number: A75250; MUID:20036896; PMID:10567266				
A:Accession: B75257				
A>Status: preliminary				
A:Molecule type: DNA				
A:Residues: 1-609 <WHI>				
A:Cross-references: UNIPROT:Q9RRC4; GB:AE002086; GB:AE000513; NID:g6460395; PIDN:AAF1210				
A:Experimental source: strain R1				
C:Genetics:				
A:Gene: DR2568				
A:Map position: 1				
C:Superfamily: Bacillus arginine-tRNA ligase				
Query Match				
Best Local Similarity 79.3%; Score 23; DB 2; Length 609;				
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;				
QY	1	VMXVAEF 7		
Db	277	IMDVSEF 283		
RESULT 62				
S33506				
protein-tyrosine kinase (EC 2.7.1.112) Cek9 - chicken (fragment)				
C:Species: Gallus gallus (chicken)				
C>Date: 22-Nov-1993 #sequence_revision 03-Aug-1995 #text_change 04-Feb-2000				
C:Accession: S33506				
R;Sajjadi, F.G.; Pasquale, E.B.				

submitted to the EMBL Data Library, December 1992				
A:Description: Five additional avian Eph-related tyrosine kinases are differentially exp				
A:Reference number: S33502				
A:Accession: S33506				
A>Status: preliminary				
A:Molecule type: DNA				
A:Residues: 1-612 <SAJ>				
A:Cross-references: EMBL:Z19060				
C:Superfamily: protein-tyrosine kinase, receptor type eph; fibronectin type III repeat hc				
C:Keywords: ATP; phosphotransferase; transmembrane protein; tyrosine-specific protein kin				
F;63-147/Domain: fibronectin type III repeat homology <3PR>				
F;245-513/Domain: protein kinase homology <KIN>				
F;253-261/Region: protein kinase ATP-binding motif				
F;536-602/Domain: SAM homology <SAM>				
Query Match				
Best Local Similarity 79.3%; Score 23; DB 2; Length 612;				
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;				
QY	1	VMXVAEF 7		
Db	321	VMIVTEF 327		
RESULT 63				
C87377				
hypothetical protein CC1031 [imported] - Caulobacter crescentus				
C:Species: Caulobacter crescentus				
C>Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004				
C:Accession: C87377				
R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.				
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon				
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.				
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001				
A>Title: Complete Genome Sequence of Caulobacter crescentus.				
A:Reference number: A87249; MUID:21173698; PMID:11259647				
A:Accession: C87377				
A>Status: preliminary				
A:Molecule type: DNA				
A:Residues: 1-618 <STO>				
A:Cross-references: UNIPROT:Q9A9F7; GB:AE005673; NID:gl3422325; PIDN:AAK23015.1; GSPDB:G				
C:Genetics:				
A:Gene: CC1031				
Query Match				
Best Local Similarity 79.3%; Score 23; DB 2; Length 618;				
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;				
QY	1	VMXVAEF 7		
Db	380	VMFVGEF 386		
RESULT 64				
JC1450				
fibroblast growth factor receptor 4 - rat				
N;Contains: protein-tyrosine kinase (EC 2.7.1.112)				
C:Species: Rattus norvegicus (Norway rat)				
C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004				
C:Accession: JC1450; PT0191				
R;Horlick, R.A.; Stack, S.L.; Cooke, G.M.				
Gene 120, 291-295, 1992				
A>Title: Cloning, expression and tissue distribution of the gene encoding rat fibroblast				
A:Reference number: JC1450; MUID:93013049; PMID:1398143				
A:Accession: JC1450				
A:Molecule type: mRNA				
A:Residues: 1-650 <HOR>				
A:Cross-references: UNIPROT:Q63709; GB:M91599; NID:g204137; PIDN:AAA41157.1; PID:g204138				
R;Lai, C.; Lemke, G.				
Neuron 6, 691-704, 1991				
A>Title: An extended family of protein-tyrosine kinase genes differentially expressed in				
A:Reference number: PT0183; MUID:91222560; PMID:2025425				
A:Accession: PT0191				

A;Molecule type: mRNA
A;Residues: 465-518 <LAI>
A;Experimental source: sciatic nerve
C;Genetics:
A;Gene: FGFR4; tyro-9
C;Function:
A;Description: receptor mediating effects of fibroblast growth factor
A;Note: expressed in normal lung; expressed in some carcinomas
C;Superfamily: basic fibroblast growth factor receptor 1; immunoglobulin homology; prote
C;Keywords: ATP; autophosphorylation; duplication; glycoprotein; growth factor receptor;
F;11-72/Domain: immunoglobulin homology <IM1>
F;110-181/Domain: immunoglobulin homology <IM2>
F;218-238/Domain: transmembrane #status predicted <TM>
F;239-650/Domain: intracellular #status predicted <INT>
F;313-598/Domain: protein kinase homology <KIN>
F;321-329/Region: protein kinase ATP-binding motif
F;104,136,157,168/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;351,368,460/Active site: Lys, Glu, Asp #status predicted
F;465,478/Binding site: magnesium (Asn, Asp) #status predicted
F;491/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status predicted

Query Match 79.3%; Score 23; DB 1; Length 650;
Best Local Similarity 57.1%; Pred. No. 4.6e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
||: |||
Db 473 VMKIADF 479

RESULT 65
E97738
DNA topoisomerase (ATP-hydrolyzing) (EC 5.99.1.3) - Rickettsia conorii (strain Malish 7)
C;Species: Rickettsia conorii
C;Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004
C;Accession: E97738
R;Ogata, H.; Audic, S.; Renesto-Audiffren, P.; Fournier, P.E.; Barbe, V.; Samson, D.; R
Science 293, 2093-2098, 2001
A;Title: Mechanisms of Evolution in Rickettsia conorii and Rickettsia prowazekii.
A;Reference number: A97700; MUID:21442074; PMID:11557893
A;Accession: E97738
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-662 <KUR>
A;Cross-references: UNIPROT:Q921M1; GB:AE006914; PIDN:AAL02847.1; PID:g15619368; GSPDB:G
C;Genetics:
A;Gene: GyrB1
C;Superfamily: DNA topoisomerase (ATP-hydrolyzing) chain B
C;Keywords: isomerase

Query Match 79.3%; Score 23; DB 2; Length 662;
Best Local Similarity 57.1%; Pred. No. 4.7e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
||: |||
Db 397 VIGIAEF 403

RESULT 66
H71676
DNA gyrase chain B (gyrB1) RP227 - Rickettsia prowazekii
C;Species: Rickettsia prowazekii
C;Date: 21-Nov-1998 #sequence_revision 21-Nov-1998 #text_change 09-Jul-2004
C;Accession: H71676
R;Andersson, S.G.E.; Zomorodipour, A.; Andersson, J.O.; Sichteritz-Ponten, T.; Alsmark, U
Nature 396, 133-140, 1998
A;Title: The genome sequence of Rickettsia prowazekii and the origin of mitochondria.
A;Reference number: A71630; MUID:99039499; PMID:9823893
A;Accession: H71676
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-662 <AND>

A;Cross-references: UNIPROT:Q9ZDU7; GB:AJ235271; GB:AJ235269; NID:g3868717; PIDN:CAA1469C
A;Experimental source: strain Madrid E
C;Genetics:
A;Gene: gyrB1; RP227
C;Superfamily: DNA topoisomerase (ATP-hydrolyzing) chain B

Query Match 79.3%; Score 23; DB 2; Length 662;
Best Local Similarity 57.1%; Pred. No. 4.7e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
||: |||
Db 397 VIGIAEF 403

RESULT 67
AB2017
two-component sensor histidine kinase all1688 [imported] - Nostoc sp. (strain PCC 7120)
C;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C;Accession: AB2017
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi,
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AB2017
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-676 <KUR>
A;Cross-references: UNIPROT:Q8YWC5; GB:BA000019; PIDN:BA078054.1; PID:g17135508; GSPDB:G
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: all1688

Query Match 79.3%; Score 23; DB 2; Length 676;
Best Local Similarity 57.1%; Pred. No. 4.8e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
||: |||
Db 516 VLNIAEF 522

RESULT 68
A35969
heparin-binding growth factor receptor K-sam precursor - human
C;Species: Homo sapiens (man)
C;Date: 09-Nov-1990 #sequence_revision 09-Nov-1990 #text_change 09-Jul-2004
C;Accession: A35969
R;Hattori, Y.; Odagiri, H.; Nakatani, H.; Miyagawa, K.; Naito, K.; Sakamoto, H.; Kato, C
Proc. Natl. Acad. Sci. U.S.A. 87, 5983-5987, 1990
A;Title: K-sam, an amplified gene in stomach cancer, is a member of the heparin-binding
A;Reference number: A35969; MUID:90332706; PMID:2377625
A;Accession: A35969
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-682 <HAT>
A;Cross-references: UNIPROT:P21802; GB:M35718; NID:g186777; PIDN:AAA36152.1; PID:g186778
C;Superfamily: basic fibroblast growth factor receptor 1; immunoglobulin homology; prote
C;Keywords: ATP; autophosphorylation; growth factor receptor; heparin binding; phosphop
F;83-144/Domain: immunoglobulin homology <IMM>
F;389-674/Domain: protein kinase homology <KIN>
F;397-405/Region: protein kinase ATP-binding motif

Query Match 79.3%; Score 23; DB 2; Length 682;
Best Local Similarity 57.1%; Pred. No. 4.8e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
||: |||
Db 549 VMKIADF 555

```
RESULT 69
S51635
fibroblast growth factor receptor 2b, keratinocyte growth factor receptor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 15-Jul-1995 #sequence_revision 21-Jul-1995 #text_change 16-Jul-1999
C;Accession: S51635; S51636; S48048; S48047
R;Savagner, P.
submitted to the EMBL Data Library, July 1994
A;Reference number: S51634
A;Accession: S51635
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-705 <SAV>
A;Cross-references: EMBL:Z35138; NID:g551271; PIDN:CAA84510.1; PID:g551272
A;Accession: S51636
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 37-172, 'T', 174-314, 'TV', 315-340, 'T', 342-705 <SA3>
A;Cross-references: EMBL:Z35139; NID:g551273; PIDN:CAA84511.1; PID:g551274
R;Savagner, P.; Valles, A.M.; Jouanneau, J.; Yamada, K.M.; Thierly, J.P.
Mol. Biol. Cell 5, 851-862, 1994
A;Title: Alternative splicing in fibroblast growth factor receptor 2 is associated with
A;Reference number: S48046; MUID:95102150; PMID:7803853
A;Accession: S48048
A;Status: preliminary; nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 37-135, 'R', 137-314, 'TV', 315-340, 'T', 342-390, 'K', 392-705 <SA2>
A;Cross-references: EMBL:Z35139
A;Accession: S48047
A;Status: preliminary; nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-135, 'R', 137-340, 'T', 342-390, 'K', 392-705 <SA4>
A;Cross-references: EMBL:Z35138
C;Superfamily: basic fibroblast growth factor receptor 1; immunoglobulin homology; protei
C;Keywords: ATP; growth factor receptor
F;57-118/Domain: immunoglobulin homology <IMM>
F;363-648/Domain: protein kinase homology <KIN>
F;371-379/Region: protein kinase ATP-binding motif

Query Match 79.3%; Score 23; DB 2; Length 705;
Best Local Similarity 57.1%; Pred. No. 5e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
Db 523 VMKIADF 529

RESULT 70
A38429
keratinocyte growth factor receptor precursor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 14-Feb-1992 #sequence_revision 14-Feb-1992 #text_change 16-Jul-1999
C;Accession: A38429
R;Miki, T.; Fleming, T.P.; Bottaro, D.P.; Rubin, J.S.; Ron, D.; Aaronson, S.A.
Science 251, 72-75, 1991
A;Title: Expression cDNA cloning of the KGF receptor by creation of a transforming autoc
A;Reference number: A38429; MUID:91095977; PMID:1846048
A;Accession: A38429
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-707 <MIK>
A;Cross-references: GB:M63503; NID:g198593; PIDN:AAA39377.1; PID:g198594
C;Superfamily: basic fibroblast growth factor receptor 1; immunoglobulin homology; protei
C;Keywords: ATP; autophosphorylation; growth factor receptor; phosphoprotein; transmembr
F;57-118/Domain: immunoglobulin homology <IMM>
F;365-650/Domain: protein kinase homology <KIN>
F;373-381/Region: protein kinase ATP-binding motif

Query Match 79.3%; Score 23; DB 2; Length 707;
Best Local Similarity 57.1%; Pred. No. 5e+02;
```

```
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
Db 525 VMKIADF 531

RESULT 71
A54846
fibroblast growth factor receptor a precursor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 28-Jul-1995 #sequence_revision 28-Jul-1995 #text_change 16-Jul-1999
C;Accession: A54846
R;Takagi, Y.; Shrivastav, S.; Miki, T.; Sakaguchi, K.
J. Biol. Chem. 269, 23743-23749, 1994
A;Title: Molecular cloning and expression of the acidic fibroblast growth factor receptor
ility and covalent attachment of heparan sulfate glycosaminoglycan to the receptors.
A;Reference number: A54846; MUID:94375484; PMID:8089146
A;Accession: A54846
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-707 <TRA>
C;Superfamily: basic fibroblast growth factor receptor 1; immunoglobulin homology; protei
C;Keywords: ATP; growth factor receptor
F;57-118/Domain: immunoglobulin homology <IMM>
F;365-650/Domain: protein kinase homology <KIN>
F;373-381/Region: protein kinase ATP-binding motif

Query Match 79.3%; Score 23; DB 2; Length 707;
Best Local Similarity 57.1%; Pred. No. 5e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
Db 525 VMKIADF 531

RESULT 72
I50128
fibroblast growth factor receptor - quail
C;Species: Coturnix coturnix (quail)
C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 09-Jul-2004
C;Accession: I50128; S42803
R;Marcelle, C.; Eichmann, A.; Halevy, O.; Breant, C.; Le Douarin, N.M.
Development 120, 683-694, 1994
A;Title: Distinct developmental expression of a new avian fibroblast growth factor recept
A;Reference number: I50128; MUID:94215505; PMID:8162862
A;Accession: I50128
A;Status: preliminary; translated from GB/EMBL/DBDJ
A;Molecule type: mRNA
A;Residues: 1-713 <MAR>
A;Cross-references: UNIPROT:Q90330; EMBL:X76885; NID:ig440139; PIDN:CAA54213.1; PID:ig44014
C;Genetics:
C;Gene: PREK
C;Superfamily: basic fibroblast growth factor receptor 1; immunoglobulin homology; protei
C;Keywords: ATP; growth factor receptor
F;176-247/Domain: immunoglobulin homology <IMM>
F;377-662/Domain: protein kinase homology <KIN>
F;385-393/Region: protein kinase ATP-binding motif

Query Match 79.3%; Score 23; DB 2; Length 713;
Best Local Similarity 57.1%; Pred. No. 5.1e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
Db 537 VMKIADF 543

RESULT 73
A56795
fibroblast growth factor receptor 1 beta-isoform - rat
C;Species: Rattus norvegicus (Norway rat)
```


C;Date: 03-Nov-1995 #sequence_revision 03-Nov-1995 #text_change 09-Jul-2004
C;Accession: A56795
R;Kim, E.G.; Kwon, H.M.; Burrow, C.R.; Ballermann, B.J.
Am. J. Physiol. 264, F66-F73, 1993
A;Title: Expression of rat fibroblast growth factor receptor 1 as three splicing variants
A;Reference number: A56795; MUID:93158789; PMID:8381605
A;Contents: Sprague-Dawley, Kidneys
A;Accession: A56795
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-729 <KIM>
A;Cross-references: UNIPROT:Q63827; GB:S54008; NID:G264804; PIDN:AAB54274.1; PID:G264805
A;Note: sequence extracted from NCBI backbone (NCBIN:124259, NCBI:P124270)
C;Superfamily: basic fibroblast growth factor receptor 1; immunoglobulin homology; prote
C;Keywords: alternative splicing; ATP; growth factor receptor
F;80-141/Domain: immunoglobulin homology <IMM>
F;383-668/Domain: protein kinase homology <KIN>
F;391-399/Region: protein kinase ATP-binding motif

Query Match 79.3%; Score 23; DB 2; Length 729;
Best Local Similarity 57.1%; Pred. No. 5.2e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
||| :||
Db 543 VMKIADF 549

RESULT 74
I49293
fibroblast growth factor receptor-1, short isoform precursor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004
C;Accession: I49293
R;Lin, Y.; Pasumarthi, K.B.; Bock, M.E.; Lytras, A.; Kardami, E.; Catini, P.A.
J. Mol. Cell. Cardiol. 26, 1449-1459, 1994
A;Title: Cloning and expression of fibroblast growth factor receptor-1 isoforms in the m
A;Reference number: I49289; MUID:95205422; PMID:7897669
A;Accession: I49293
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-733 <RES>
A;Cross-references: UNIPROT:Q60830; EMBL:U23445; NID:G733537; PIDN:AAC52183.1; PID:G7335
C;Superfamily: basic fibroblast growth factor receptor 1; immunoglobulin homology; prote
C;Keywords: ATP; growth factor receptor
F;181-254/Domain: immunoglobulin homology <IMM>
F;387-672/Domain: protein kinase homology <KIN>
F;395-403/Region: protein kinase ATP-binding motif

Query Match 79.3%; Score 23; DB 2; Length 733;
Best Local Similarity 57.1%; Pred. No. 5.2e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
||| :||
Db 547 VMKIADF 553

RESULT 75
S41050
fibroblast growth factor receptor-2 - eastern newt
C;Species: Notophthalmus viridescens, Triturus viridescens (eastern newt)
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 09-Jul-2004
C;Accession: S41050
R;Poulin, M.L.; Chiu, I.M.
Biochim. Biophys. Acta 1220, 209-211, 1994
A;Title: Nucleotide sequences of two new (Notophthalmus viridescens) fibroblast growth
A;Reference number: S41050; MUID:94146117; PMID:8312364
A;Accession: S41050
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-748 <POU>
A;Cross-references: UNIPROT:Q91147

C;Superfamily: basic fibroblast growth factor receptor 1; immunoglobulin homology; prote
F;101-162/Domain: immunoglobulin homology <IMM>
F;406-691/Domain: protein kinase homology <KIN>

Query Match 79.3%; Score 23; DB 2; Length 748;
Best Local Similarity 57.1%; Pred. No. 5.3e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
||| :||
Db 566 VMKIADF 572

RESULT 76
S41051
fibroblast growth factor receptor-2 - eastern newt
C;Species: Notophthalmus viridescens, Triturus viridescens (eastern newt)
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 09-Jul-2004
C;Accession: S41051
R;Poulin, M.L.; Chiu, I.M.
Biochim. Biophys. Acta 1220, 209-211, 1994
A;Title: Nucleotide sequences of two new (Notophthalmus viridescens) fibroblast growth
A;Reference number: S41050; MUID:94146117; PMID:8312364
A;Accession: S41051
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-750 <POU>
A;Cross-references: UNIPROT:Q91150
C;Superfamily: basic fibroblast growth factor receptor 1; immunoglobulin homology; prote
F;101-162/Domain: immunoglobulin homology <IMM>
F;408-693/Domain: protein kinase homology <KIN>

Query Match 79.3%; Score 23; DB 2; Length 750;
Best Local Similarity 57.1%; Pred. No. 5.3e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
||| :||
Db 568 VMKIADF 574

RESULT 77
S16236
fibroblast growth factor receptor precursor - human
N;Alternate names: bek-related FGF receptor
N;Contains: protein-tyrosine kinase (EC 2.7.1.112)
C;Species: Homo sapiens (man)
C;Date: 16-Sep-1992 #sequence_revision 16-Sep-1992 #text_change 09-Jul-2004
C;Accession: S16236
R;Seno, M.; Sasada, R.; Watanabe, T.; Ishimaru, K.; Igarashi, K.
Biochim. Biophys. Acta 1089, 244-246, 1991
A;Title: Two cDNAs encoding novel human FGF receptor.
A;Reference number: S16236; MUID:91274356; PMID:1647213
A;Accession: S16236
A;Molecule type: mRNA
A;Residues: 1-769 <SEN>
A;Cross-references: UNIPROT:P21802; EMBL:X56191; NID:G29431; PIDN:CAA39654.1; PID:G29432
C;Superfamily: basic fibroblast growth factor receptor 1; immunoglobulin homology; prote
C;Keywords: ATP; autophosphorylation; duplication; glycoprotein; growth factor receptor;
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-769/Product: fibroblast growth factor receptor #status predicted <MAT>
F;22-378/Domain: extracellular #status predicted <EXT>
F;132-138/Region: acidic
F;172-233/Domain: immunoglobulin homology <IMM>
F;379-399/Domain: transmembrane #status predicted <TMW>
F;400-769/Domain: intracellular #status predicted <INT>
F;480-765/Domain: protein kinase homology <KIN>
F;488-496/Region: protein kinase ATP-binding motif
F;62-107,179-231,278-340/Disulfide bonds: #status predicted
F;83,123,228,241,265,297,318,329,350/Asp #status predicted
F;518,535,627/Active site: Lys, Glu, Asp #status predicted

Query Match 79.3%; Score 23; DB 2; Length 769;

Best Local Similarity 57.1%; Pred. No. 5.4e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMKVAEF 7
|| :||
Db 640 VMKIADF 646

RESULT 78

B71931
DNA gyrase chain B - Helicobacter pylori (strain J99)
C:Species: Helicobacter pylori
A:Variety: strain J99
C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 09-Jul-2004
C:Accession: B71931
R;Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.;
Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.;
Nature 397, 176-180, 1999
A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric path
A:Reference number: A71800; MUID:99120557; PMID:9923682
A:Accession: B71931
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-773 <ARN>
A:Cross-references: UNIPROT:Q9LX3; GB:AE001479; GB:AE001439; NID:94154979; PIDN:AAD0603
A:Experimental source: strain J99
C:Genetics:
A:Gene: gyrB
C:Superfamily: DNA topoisomerase (ATP-hydrolyzing) chain B

Query Match 79.3%; Score 23; DB 2; Length 773;
Best Local Similarity 71.4%; Pred. No. 5.5e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 VMKVAEF 7
|| :||
Db 176 VMEVVEF 182

RESULT 79

E64582
DNA gyrase, sub B - Helicobacter pylori (strain 26695)
C:Species: Helicobacter pylori
C:Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 09-Jul-2004
C:Accession: E64582
R;Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.;
Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKenne
son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L.
Nature 388, 539-547, 1997
A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.
A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.
A:Reference number: A64520; MUID:97394467; PMID:9252185
A:Accession: E64582
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-773 <TOM>
A:Cross-references: UNIPROT:P55992; GB:AE000564; GB:AE000511; NID:92313602; PIDN:AAD0756
C:Superfamily: DNA topoisomerase (ATP-hydrolyzing) chain B

Query Match 79.3%; Score 23; DB 2; Length 773;
Best Local Similarity 71.4%; Pred. No. 5.5e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 VMKVAEF 7
|| :||
Db 176 VMEVVEF 182

RESULT 80

A39627
protein-tyrosine kinase (EC 2.7.1.112) homolog - fruit fly (Drosophila melanogaster) (fr
N;Alternate names: fibroblast growth factor receptor
C:Species: Drosophila melanogaster

C:Date: 24-Jan-1992 #sequence_revision 24-Jan-1992 #text_change 04-Feb-2000
C:Accession: A39627; S16679
R;Glazer, L.; Shilo, B.Z.
Genes Dev. 5, 697-705, 1991
A:Title: The Drosophila PGF-R homolog is expressed in the embryonic tracheal system and i
A:Reference number: A39627; MUID:91184623; PMID:1849109
A:Accession: A39627
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-790 <GLA>
A:Cross-references: EMBL:X57746; NID:g7965; PIDN:CAA40912.1; PID:g7966
C:Genetics:
A:Gene: FlyBase:bt1
A:Cross-references: FlyBase:FBgn0005592
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo
C:Keywords: ATP; growth factor receptor; phosphotransferase; tyrosine-specific protein ki
F;449-741/Domain: protein kinase homology <KIN>
F;457-465/Region: protein kinase ATP-binding motif

Query Match 79.3%; Score 23; DB 2; Length 790;
Best Local Similarity 57.1%; Pred. No. 5.6e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMKVAEF 7
|| :||
Db 616 VMKIADF 622

RESULT 81

T41573
hypothetical protein SPCC736.16 - fission yeast (Schizosaccharomyces pombe)
N;Alternate names: hypothetical protein SPC594.01
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 17-Nov-2000 #text_change 09-Jul-2004
C:Accession: T41573; T41445
R;Wood, V.; Rajandream, M.A.; Barrell, B.G.; Murphy, L.; Harris, D.
submitted to the EMBL Data Library, May 1998
A:Reference number: Z21991
A:Accession: T41573
A:Status: translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-231 <WOO>
A:Cross-references: UNIPROT:O74504; EMBL:AL023705; NID:g3169070; PIDN:CAA19280.1; PID:g31
A:Experimental source: strain 972h-; cosmid c736
R;Rieger, M.; Lyne, M.; Rajandream, M.A.; Barrell, B.G.
submitted to the EMBL Data Library, September 1998
A:Reference number: Z21994
A:Accession: T41445
A:Status: translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 95-791 <RIE>
A:Cross-references: EMBL:AL031523; NID:g3560217; PIDN:CAA20660.1; PID:g3560218; GSPDB:GN
A:Experimental source: strain 972h-; cosmid c594
C:Genetics:
A:Gene: SPDB:SPCC736.16; SPDB:SPCC594.01
A:Map position: 3
A:Introns: 101/1; 131/2

Query Match 79.3%; Score 23; DB 2; Length 791;
Best Local Similarity 71.4%; Pred. No. 5.6e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 VMKVAEF 7
|| :||
Db 761 VMFVVEF 767

RESULT 82

S38579
fibroblast growth factor receptor 3 - Iberian ribbed newt (fragment)
C:Species: Pleurodeles waltl (Iberian ribbed newt)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004
C:Accession: S38579

R;Shi, D.L.; Fromentoux, V.; Launay, C.; Umbhauer, M.; Boucaut, J.C.
submitted to the EMBL Data Library, November 1993
A;Description: Expression of FGFR-3 in amphibian embryos.

A;Reference number: S38579

A;Accession: S38579

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-797 <SHI>

A;Cross-references: UNIPROT:Q91287; EMBL:X75603

C;Superfamily: basic fibroblast growth factor receptor 1; immunoglobulin homology; prote

C;Keywords: ATP; growth factor receptor

F;258-331/Domain: immunoglobulin homology <IMM>

F;456-741/Domain: protein kinase homology <KIN>

F;464-472/Region: protein kinase ATP-binding motif

Query Match 79.3%; Score 23; DB 2; Length 797;

Best Local Similarity 57.1%; Pred. No. 5.6e+02;

Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VMXVAEF 7

Db 616 VMKIADF 622

RESULT 83

S18209

fibroblast growth factor receptor 4 precursor (clone 61) - mouse

N;Alternate names: tyrosine kinase Mpk-11

C;Keywords: ATP; autophosphorylation; duplication; glycoprotein; growth factor receptor;

C;Species: Mus musculus (house mouse)

C;Date: 07-Oct-1994 #sequence_revision 07-Oct-1994 #text_change 09-Jul-2004

C;Accession: S18209; S26751; S30497

R;Stark, K.L.; McMahon, J.A.; McMahon, A.P.

Development 113, 641-651, 1991

A;Title: FGFR-4, a new member of the fibroblast growth factor receptor family, expressed

A;Reference number: S18209; MUID:92146274; PMID:1723680

A;Accession: S18209

A;Molecule type: mRNA

A;Residues: 1-799 <STAL>

A;Cross-references: UNIPROT:Q03142; EMBL:X59927

R;Stark, K.L.

submitted to the EMBL Data Library, May 1991

A;Reference number: S26751

A;Accession: S26751

A;Molecule type: mRNA

A;Residues: 1-485, 'QVVRAEAFG', 486-799 <STA2>

A;Cross-references: EMBL:X59927; NID:G50968; PIDN:CAA42551.1; PID:G50969

R;Gliardi-Hebenstreit, P.; Nieto, M.A.; Frain, M.; Mattei, M.G.; Chestier, A.; Wilkinson

Oncogene 7, 2499-2506, 1992

A;Title: An Eph-related receptor protein tyrosine kinase gene segmentally expressed in b

A;Reference number: S30496; MUID:93096484; PMID:1281307

A;Accession: S30497

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 611-667 <GIL>

A;Cross-references: EMBL:X57236; NID:G53187; PIDN:CAA40512.1; PID:G53188

C;Superfamily: basic fibroblast growth factor receptor 1; immunoglobulin homology; prote

C;Keywords: ATP; duplication; glycoprotein; growth factor receptor; phosphotransferase;

F;1-18/Domain: signal sequence #status predicted <SIG>

F;1-799/Product: fibroblast growth factor receptor 4 #status predicted <MAT>

F;19-366/Domain: extracellular #status predicted <EXT>

F;162-223/Domain: immunoglobulin homology <IMM>

F;367-387/Domain: transmembrane #status predicted <TMW>

F;388-799/Domain: intracellular #status predicted <INT>

F;482-747/Domain: protein kinase homology <KIN>

F;470-478/Region: protein kinase ATP-binding motif

F;54-98,169-221,268-330/Disulfide bonds: #status predicted

F;500,517,609/Active site: Lys, Glu, Asp #status predicted

Query Match

Best Local Similarity 79.3%; Score 23; DB 2; Length 799;

Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VMXVAEF 7

Db 622 VMKIADF 628

RESULT 84

TVHU2F

fibroblast growth factor receptor flg-2 precursor - human

N;Contains: protein-tyrosine kinase (EC 2.7.1.112) flg-2

C;Species: Homo sapiens (man)

C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 16-Jun-2000

C;Accession: A60350; S21843

R;Avivi, A.; Zimmer, Y.; Yavon, A.; Yarden, Y.; Givol, D.

Oncogene 6, 1089-1092, 1991

A;Title: Flg-2, a new member of the family of fibroblast growth factor receptors.

A;Reference number: A60350; MUID:91296390; PMID:1648703

A;Accession: A60350

A;Molecule type: mRNA

A;Residues: 1-800 <AVI>

A;Cross-references: EMBL:X58255; NID:G31382; PIDN:CAA41209.1; PID:G31383

A;Experimental source: Keratinocytes

C;Comment: This may be a receptor for keratinocyte growth factor.

C;Genetics:

A;Gene: GDB:FGFR2; JWS; CFD1; KGF; FLG2

A;Cross-references: GDB:127273; OMIM:176943

A;Map position: 10q25.3-10q26

C;Superfamily: basic fibroblast growth factor receptor 1; immunoglobulin homology; prote

C;Keywords: ATP; autophosphorylation; duplication; glycoprotein; growth factor receptor;

F;1-21/Domain: signal sequence #status predicted <SIG>

F;22-800/Product: fibroblast growth factor receptor flg-2 #status predicted <MAT>

F;22-369/Domain: extracellular #status predicted <EXT>

F;131-137/Region: acidic

F;262-335/Domain: immunoglobulin homology <IMM>

F;370-390/Domain: transmembrane #status predicted <TMW>

F;391-800/Domain: intracellular #status predicted <INT>

F;464-749/Domain: protein kinase homology <KIN>

F;472-480/Region: protein kinase ATP-binding motif

F;59-107,170-222,269-333/Disulfide bonds: #status predicted

F;56,219,256,288,309,322/Binding site: carboxylate (Asn) (covalent) #status predicted

F;502,519,611/Active site: Lys, Glu, Asp #status predicted

F;616,629/Binding site: magnesium (Asn, Asp) #status predicted

F;642/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status predicted

Query Match 79.3%; Score 23; DB 1; Length 800;

Best Local Similarity 57.1%; Pred. No. 5.7e+02;

Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VMXVAEF 7

Db 624 VMKIADF 630

RESULT 85

A48991

heparin-binding growth factor receptor - mouse

N;Alternate names: HBGF receptor

N;Contains: protein-tyrosine kinase (EC 2.7.1.112)

C;Species: Mus musculus (house mouse)

C;Date: 19-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004

C;Accession: A48991

R;Kato, O.; Hattori, Y.; Sasaki, H.; Sakamoto, H.; Fujimoto, K.; Fujii, T.; Sugimura, T.

Cancer Res. 53, 1136-1141, 1993

A;Title: Isolation of the complementary DNA encoding a mouse heparin-binding growth fact

A;Reference number: A48991; MUID:93177694; PMID:8382556

A;Accession: A48991

A;Status: preliminary

A;Molecule type: nucleic acid

A;Residues: 1-800 <KAT>

A;Cross-references: UNIPROT:Q61851; GB:S56291; NID:G298329; PIDN:AAB25535.1; PID:G298330

A;Experimental source: brain

A;Note: sequence extracted from NCBI backbone (NCBI:126536, NCBI:126537)

C;Superfamily: basic fibroblast growth factor receptor 1; immunoglobulin homology; prote

C;Keywords: ATP; growth factor receptor; heparin binding; phosphotransferase; tyrosine-s

F;262-335/Domain: immunoglobulin homology <IMM>
F;464-749/Domain: protein kinase homology <KIN>
F;472-480/Region: protein kinase ATP-binding motif

Query Match 79.3%; Score 23; DB 2; Length 800;
Best Local Similarity 57.1%; Pred. No. 5.7e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VMKVAEF 7
|| :||
Db 624 VMKIADF 630

RESULT 86
I55363
fibroblast growth factor receptor 3 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004
C;Accession: I55363; B53627
R;Ornitz, D.M.; Leder, P.
J. Biol. Chem. 267, 16305-16311, 1992
A;Title: Ligand specificity and heparin dependence of fibroblast growth factor receptors
A;Reference number: I55363; MUID:92355591; PMID:1379594
A;Accession: I55363
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-801 <RES>
A;Cross-references: UNIPROT:Q61851; GB:M81342; NID:gi199144; PIDN:AAA39535.1; PID:gi199145
R;Chellaiiah, A.T.; McEwen, D.G.; Werner, S.; Xu, J.; Ornitz, D.M.
J. Biol. Chem. 269, 11620-11627, 1994
A;Title: Fibroblast growth factor receptor (FGFR) 3. Alternative splicing in immunoglobulin
A;Reference number: A53627; MUID:94209351; PMID:7512569
A;Accession: B53627
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 242-364 <CHB>
A;Cross-references: GB:L26492
C;Genetics:

A;Gene: mPR3
A;Introns: 304/3; 353/1
C;Superfamily: basic fibroblast growth factor receptor 1; immunoglobulin homology; protein
C;Keywords: Atg; growth factor receptor
F;262-335/Domain: immunoglobulin homology <IMM>
F;464-749/Domain: protein kinase homology <KIN>
F;472-480/Region: protein kinase ATP-binding motif

Query Match 79.3%; Score 23; DB 2; Length 801;
Best Local Similarity 57.1%; Pred. No. 5.7e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VMKVAEF 7
|| :||
Db 624 VMKIADF 630

RESULT 87
TVHUF4
fibroblast growth factor receptor 4 precursor - human
N;Alternate names: protein-tyrosine kinase tkf
N;Contains: protein-tyrosine kinase (EC 2.7.1.112)
C;Species: Homo sapiens (man)
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
C;Accession: S15345; A46615; A41598; D38269
R;Partanen, J.; Maekelae, T.P.; Eroila, E.; Korhonen, J.; Hirvonen, H.; Claesson-Welsh, E.
EMBO J. 10, 1347-1354, 1991
A;Title: FGFR-4, a novel acidic fibroblast growth factor receptor with a distinct expression
A;Reference number: S15345; MUID:91224085; PMID:1709094
A;Accession: S15345
A;Molecule type: mRNA
A;Residues: 1-802 <PAR>
A;Cross-references: UNIPROT:P22455; EMBL:X57205; NID:g31371; PIDN:CAA04090.1; PID:g31372
A;Note: binds acidic but not basic fibroblast growth factor with high affinity
R;Ron, D.; Reich, R.; Chedid, M.; Lengel, C.; Cohen, O.E.; Chan, A.M.; Neufeld, G.; Miki

J. Biol. Chem. 268, 5388-5394, 1993
A;Title: Fibroblast growth factor receptor 4 is a high affinity receptor for both acidic and basic fibroblast growth factors with high affinity
A;Reference number: A46615; MUID:93194827; PMID:7680645
A;Accession: A46615
A;Molecule type: mRNA
A;Status: nucleic acid sequence not shown; not compared with conceptual translation
A;Residues: 1-296, 'D', 298-802 <RON>
A;Experimental source: mammary epithelial cell line B5/589
A;Note: sequence extracted from NCBI backbone (NCBIP:127650)
R;Holtrich, U.; Braeuninger, A.; Strebhardt, K.; Ruebsamen-Waigmann, H.
Proc. Natl. Acad. Sci. U.S.A. 88, 10411-10415, 1991
A;Title: Two additional protein-tyrosine kinases expressed in human lung: fourth member of the fibroblast growth factor receptor family
A;Reference number: S19025; MUID:92073297; PMID:1720539
A;Accession: A41598
A;Status: nucleic acid sequence not shown; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 399-534, 'M', 536-799, 'SG', 800-802 <HOL>
A;Experimental source: lung
R;Partanen, J.; Maekelae, T.P.; Alitalo, R.; Lehtvaeslahti, H.; Alitalo, K.
Proc. Natl. Acad. Sci. U.S.A. 87, 8913-8917, 1990
A;Title: Putative tyrosine kinases expressed in K-562 human leukemia cells.
A;Reference number: A38268; MUID:91062389; PMID:2247464
A;Accession: D38269
A;Molecule type: mRNA
A;Residues: 614-670 <PA2>
A;Cross-references: GB:M37781
A;Experimental source: K-562 leukemia cell line
C;Genetics:
A;Gene: GDB:FGFR4
A;Cross-references: GDB:127929; OMIM:134935
A;Map position: 5q33.2-5qter
C;Function:
A;Description: receptor mediating effects of fibroblast growth factor
A;Note: expressed in normal lung; expressed in some carcinomas
C;Superfamily: basic fibroblast growth factor receptor 1; immunoglobulin homology; protein
C;Keywords: ATP; autophosphorylation; duplication; glycoprotein; growth factor receptor;
F;1-24/Domain: signal sequence #status predicted <SIG>
F;25-802/Product: fibroblast growth factor receptor 4 #status predicted <MAT>
F;25-369/Domain: extracellular #status predicted <EXT>
F;50-103/Domain: immunoglobulin homology <IM1>
F;165-226/Domain: immunoglobulin homology <IM2>
F;264-335/Domain: immunoglobulin homology <IM3>
F;370-390/Domain: transmembrane #status predicted <TMW>
F;391-802/Domain: intracellular #status predicted <INT>
F;465-750/Domain: protein kinase homology <KIN>
F;473-481/Region: protein kinase ATP-binding motif
F;57-101,172-224,271-333/Dissulfide bonds: #status predicted
F;112,258,290,311,322/Binding site: carboxylate (Asn) (covalent) #status predicted
F;503,520,612/Active site: Lys, Glu, Asp #status predicted
F;617,630/Binding site: magnesium (Asn, Asp) #status predicted
F;643/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status predicted

Query Match 79.3%; Score 23; DB 1; Length 802;
Best Local Similarity 57.1%; Pred. No. 5.7e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VMKVAEF 7
|| :||
Db 625 VMKIADF 631

RESULT 88
TVHUF3
fibroblast growth factor receptor 3 precursor - human
N;Contains: protein-tyrosine kinase (EC 2.7.1.112)
C;Species: Homo sapiens (man)
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
C;Accession: A38576; A55273; E38269; I51880
R;Keegan, K.; Johnson, D.E.; Williams, L.T.; Hayman, M.J.
Proc. Natl. Acad. Sci. U.S.A. 88, 1095-1099, 1991
A;Title: Isolation of an additional member of the fibroblast growth factor receptor family
A;Reference number: A38576; MUID:91142118; PMID:1847508

A;Accession: A38576
A;Molecule type: mRNA
A;Residues: 1-806 <KEE>
A;Cross-references: UNIPROT:P22607; GB:M58051; NID:G182568; PIDN:AAA52450.1; PID:G182569
R;Thompson, L.M.; Plummer, S.; Schalling, M.; Altherr, M.R.; Guella, J.F.; Houseman, D.E
Genomics 11, 1133-1142, 1991
A;Title: A gene encoding a fibroblast growth factor receptor isolated from the Huntington
A;Reference number: A55273; MUID:92147110; PMID:1664411
A;Accession: A55273
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 76-394, 'V', 396-806 <THO>
A;Cross-references: GB:M64347; NID:G182564; PIDN:AAA58470.1; PID:G182565
A;Note: sequence extracted from NCBI backbone (NCBIP:80296)
R;Partanen, J.; Mäkelä, T.P.; Alitalo, R.; Leivaeslaiho, H.; Alitalo, K.
Proc. Natl. Acad. Sci. U.S.A. 87, 8913-8917, 1990
A;Title: Putative tyrosine kinases expressed in K-562 human leukemia cells.
A;Reference number: A38468; MUID:91062389; PMID:2247464
A;Accession: E38269
A;Molecule type: mRNA
A;Residues: 619-675 <PAR>
A;Cross-references: GB:M37782
R;Bellus, G.A.; Hefferon, T.W.; Ortiz de Luna, R.I.; Hecht, J.T.; Horton, W.A.; Machado,
Am. J. Hum. Genet. 56, 368-373, 1995
A;Title: Achondroplasia is defined by recurrent G380R mutations of FGFR3.
A;Reference number: I51880; MUID:95150025; PMID:7847369
A;Accession: I51880
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 361-379, 'R', 381-415 <RES>
A;Cross-references: GB:S76733; NID:G914201; PIDN:AAB33323.1; PID:G914202
A;Note: this sequence represents a mutant form associated with achondroplasia
C;Genetics:
A;Gene: GDB:FGFR3
A;Cross-references: GDB:I27526; OMIM:100800; OMIM:134934
A;Map position: 4p16.3-4p16.3
C;Function:
A;Description: receptor for both acidic and basic fibroblast growth factors
C;Superfamily: basic fibroblast growth factor receptor 1; immunoglobulin homology; protei
C;Keywords: ATP; autophosphorylation; duplication; glycoprotein; growth factor receptor;
F;1-22/Domain: signal sequence #status predicted <SIG>
F;23-806/Product: fibroblast growth factor receptor 3 #status predicted <MAT>
F;23-375/Domain: extracellular #status predicted <EXT>
F;133-139/Region: acidic
F;268-341/Domain: immunoglobulin homology <IMM>
F;376-396/Domain: transmembrane #status predicted <TMW>
F;397-806/Domain: intracellular #status predicted <INT>
F;470-755/Domain: protein kinase homology <KIN>
F;478-486/Region: protein kinase ATP-binding motif
F;61-109,176-228,275-339/Disulfide bonds: #status predicted
F;98,225,262,294,315,328/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;508,525,617/Active site: Lys, Glu, Asp #status predicted
F;622,635/Binding site: magnesium (Asn, Asp) #status predicted
F;648/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status predicted

Query Match 79.3%; Score 23; DB 1; Length 806;
Best Local Similarity 57.1%; Pred. No. 5.7e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMKVAEF 7
|| :||
Db 630 VKMIADF 636

RESULT 89
A35963
protein-tyrosine kinase (EC 2.7.1.112) cek2 precursor - chicken
C;Species: Gallus gallus (chicken)
C;Date: 09-Nov-1990 #sequence_revision 09-Nov-1990 #text_change 09-Jul-2004
C;Accession: A35963
R;Pasquale, E.B.
Proc. Natl. Acad. Sci. U.S.A. 87, 5812-5816, 1990
A;Title: A distinctive family of embryonic protein-tyrosine kinase receptors.

A;Reference number: A35963; MUID:90332672; PMID:2165604
A;Accession: A35963
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-806 <PAS>
A;Cross-references: UNIPROT:P18460; GB:M35195; NID:G211442; PIDN:AAA48664.1; PID:G211443
C;Genetics:
A;Gene: cek2
C;Superfamily: basic fibroblast growth factor receptor 1; immunoglobulin homology; protei
C;Keywords: ATP; autophosphorylation; duplication; glycoprotein; growth factor receptor;
F;1-23/Domain: signal sequence #status predicted <SIG>
F;24-806/Product: protein-tyrosine kinase cek2 #status predicted <MAT>
F;24-368/Domain: extracellular #status predicted <EXT>
F;131-138/Region: acidic
F;262-335/Domain: immunoglobulin homology <IMM>
F;369-389/Domain: transmembrane #status predicted <TMW>
F;390-806/Domain: intracellular #status predicted <INT>
F;464-749/Domain: protein kinase homology <KIN>
F;472-480/Region: protein kinase ATP-binding motif
F;61-107,170-222,269-333/Disulfide bonds: #status predicted
F;96,219,256,288,309,322/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;502,519,611/Active site: Lys, Glu, Asp #status predicted
F;616,629/Binding site: magnesium (Asn, Asp) #status predicted
F;642/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status predicted

Query Match 79.3%; Score 23; DB 2; Length 806;
Best Local Similarity 57.1%; Pred. No. 5.7e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMKVAEF 7
|| :||
Db 624 VKMIADF 630

RESULT 90
A39752
fibroblast growth factor receptor A1 precursor - African clawed frog
N;Contains: fibroblast growth factor receptor A1, short splice form; protein-tyrosine ki
C;Species: Xenopus laevis (African clawed frog)
C;Date: 07-Feb-1992 #sequence_revision 19-Jan-1996 #text_change 09-Jul-2004
C;Accession: A39752; B39752
R;Priesel, R.; Dawid, I.B.
Mol. Cell. Biol. 11, 2481-2488, 1991
A;Title: cDNA cloning and developmental expression of fibroblast growth factor receptors
A;Reference number: A39752; MUID:91203867; PMID:1850097
A;Accession: A39752
A;Molecule type: mRNA
A;Residues: 1-814 <FRI>
A;Cross-references: UNIPROT:Q91897; GB:M55163; NID:G214893; PIDN:AAA49990.1; PID:G214894
A;Accession: B39752
A;Molecule type: mRNA
A;Residues: 1-30,119-814 <PR2>
A;Cross-references: GB:M55163
C;Superfamily: basic fibroblast growth factor receptor 1; immunoglobulin homology; protei
C;Keywords: alternative splicing; ATP; autophosphorylation; duplication; glycoprotein; g
protein kinase
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-814/Product: fibroblast growth factor receptor A1, long splice form #status experim
F;22-30,119-814/Product: fibroblast growth factor receptor A1, short splice form #status
F;47-102/Domain: immunoglobulin homology <IM1>
F;125-132/Region: acidic
F;167-228/Domain: immunoglobulin homology <IM2>
F;266-339/Domain: immunoglobulin homology <IM3>
F;373-393/Domain: transmembrane #status predicted <TMW>
F;394-814/Domain: intracellular #status predicted <INT>
F;472-757/Domain: protein kinase homology <KIN>
F;480-488/Region: protein kinase ATP-binding motif
F;54-100,174-226,273-337/Disulfide bonds: #status predicted
F;76,116,133,177,223,236,260,292,313,326/Binding site: carbohydrate (Asn) (covalent) #sta
F;510,527,619/Active site: Lys, Glu, Asp #status predicted
F;624,637/Binding site: magnesium (Asn, Asp) #status predicted
F;650/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status predictor

Query Match 79.3%; Score 23; DB 1; Length 814;
 Best Local Similarity 57.1%; Pred. No. 5.8e+02;
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
 ||| :|||
 Db 632 VMKIADF 638

RESULT 91
 A49151
 fibroblast growth factor receptor 1 - Iberian ribbed newt
 C:Species: Pleurodeles wallii (Iberian ribbed newt)
 C>Date: 19-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
 C:Accession: A49151
 R:Shi, D.L.; Feige, J.J.; Riou, J.F.; DeSimone, D.W.; Boucaut, J.C.
 Development 116, 261-273, 1992
 A:Title: Differential expression and regulation of two distinct fibroblast growth factor
 A:Reference number: A49151; MUID:93130775; PMID:1483392
 A:Accession: A49151
 A:Status: preliminary
 A:Molecule type: nucleic acid
 A:Residues: 1-816 <SHI>
 A:Cross-references: UNIPROT:Q91285
 A:Note: sequence extracted from NCBI backbone (NCBIN:122562, NCBIP:122597)
 C:Superfamily: basic fibroblast growth factor receptor 1; immunoglobulin homology; prote
 C:Keywords: ATP; growth factor receptor
 F:169-230/Domain: immunoglobulin homology <IMM>
 F:472-757/Domain: protein kinase homology <KIN>
 F:480-488/Region: protein kinase ATP-binding motif

Query Match 79.3%; Score 23; DB 2; Length 816;
 Best Local Similarity 57.1%; Pred. No. 5.8e+02;
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
 ||| :|||
 Db 632 VMKIADF 638

RESULT 92
 JC4058
 fibroblast growth factor receptor-4 precursor - African clawed frog
 C:Species: Xenopus laevis (African clawed frog)
 C>Date: 29-Jun-1995 #sequence_revision 14-Jul-1995 #text_change 09-Jul-2004
 C:Accession: JC4058
 R:Shiozaki, C.; Tashiro, K.; Asano-Miyoshi, M.; Saigo, K.; Emori, Y.; Shiohawa, K.
 Gene 152, 215-219, 1995
 A:Title: Cloning of cDNA and genomic DNA encoding fibroblast growth factor rece ptor-4 c
 A:Reference number: JC4058; MUID:95137391; PMID:7835703
 A:Accession: JC4058
 A:Molecule type: mRNA
 A:Residues: 1-818 <SHI>
 A:Cross-references: UNIPROT:Q91742; DBJ:D31761; NID:g809527; PIDN:BA06539.1; PID:g8095
 C:Genetics:
 A:Introns: 43/1; 125/1; 158/1; 213/1; 254/1; 318/1; 364/1; 429/1; 478/1; 556/1; 619/1; 6
 C:Superfamily: basic fibroblast growth factor receptor 1; immunoglobulin homology; prote
 C:Keywords: ATP; growth factor receptor; transmembrane protein
 F:1-26/Domain: signal sequence #status predicted <SIG>
 F:27-818/Product: fibroblast growth factor receptor-4 #status predicted <MAT>
 F:56-110/Domain: immunoglobulin homology <IM1>
 F:132-137/Domain: acidic #status predicted <ADI>
 F:177-238/Domain: immunoglobulin homology <IM2>
 F:276-347/Domain: immunoglobulin homology <IM3>
 F:382-402/Domain: transmembrane #status predicted <TMM>
 F:477-762/Domain: protein kinase homology <KIN>
 F:485-493/Region: protein kinase ATP-binding motif

Query Match 79.3%; Score 23; DB 2; Length 818;
 Best Local Similarity 57.1%; Pred. No. 5.8e+02;
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
 ||| :|||
 Db 637 VMKIADF 643

RESULT 93
 TVCHFG
 fibroblast growth factor receptor 1 precursor - chicken
 N:Alternate names: basic fibroblast growth factor receptor
 N:Contains: protein-tyrosine kinase (EC 2.7.1.112) cck1
 C:Species: Gallus gallus (chicken)
 C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
 C:Accession: A41345; A33908
 R:Lee, P.L.; Johnson, D.E.; Cousens, L.S.; Fried, V.A.; Williams, L.T.
 Science 245, 57-60, 1989
 A:Title: Purification and complementary DNA cloning of a receptor for basic fibroblast g
 A:Reference number: A41345; MUID:89298406; PMID:2544996
 A:Accession: A41345
 A:Status: nucleic acid sequence not shown; not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 1-819 <DEE>
 A:Cross-references: UNIPROT:P21804
 A:Note: part of the sequence was confirmed by protein sequencing
 R:Pasquale, E.B.; Singer, S.J.
 Proc. Natl. Acad. Sci. U.S.A. 86, 5449-5453, 1989
 A:Title: Identification of a developmentally regulated protein-tyrosine kinase by using e
 A:Reference number: A33908; MUID:89315814; PMID:2473471
 A:Accession: A33908
 A:Molecule type: mRNA
 A:Residues: 1-89, 'A', '91-685', 'M', 687-819 <PAS>
 A:Cross-references: GB:M24637
 A:Note: this protein is expressed in embryonic tissues and, at low levels, in adult brain
 C:Genetics:
 A:Gene: cck1
 C:Superfamily: basic fibroblast growth factor receptor 1; immunoglobulin homology; protei
 C:Keywords: ATP; autophosphorylation; duplication; glycoprotein; growth factor receptor;
 F:1-21/Domain: signal sequence #status predicted <SIG>
 F:22-819/Product: fibroblast growth factor receptor 1 #status predicted <MAT>
 F:22-374/Domain: extracellular #status predicted <EXT>
 F:125-132/Region: acidic
 F:169-230/Domain: immunoglobulin homology <IMM>
 F:375-395/Domain: transmembrane #status predicted <TMM>
 F:396-819/Domain: intracellular #status predicted <INT>
 F:474-759/Domain: protein kinase homology <KIN>
 F:482-490/Region: protein kinase ATP-binding motif
 F:54-100,176-228,275-339/Disulfide bonds: #status predicted
 F:76,116,225,238,262,294,315,328/Binding site: carbohydrate
 F:512,529,621/Active site: Lys, Glu, Asp #status predicted
 F:626,639/Binding site: magnesium (Asn, Asp) #status predicted
 F:652/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status predicted

Query Match 79.3%; Score 23; DB 1; Length 819;
 Best Local Similarity 57.1%; Pred. No. 5.8e+02;
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
 ||| :|||
 Db 634 VMKIADF 640

RESULT 94
 S17295
 fibroblast growth factor receptor - mouse
 C:Species: Mus musculus (house mouse)
 C>Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 09-Jul-2004
 C:Accession: S17295
 R:Raz, V.; Kelman, Z.; Avivi, A.; Neufeld, G.; Givol, D.; Yarden, Y.
 Oncogene 6, 753-760, 1991
 A:Title: PCR-based identification of new receptors: molecular cloning of a receptor for f
 A:Reference number: S17295; MUID:91270892; PMID:1711190
 A:Accession: S17295
 A:Status: preliminary
 A:Molecule type: mRNA

A;Residues: 1-820 <RAZ>
 A;Cross-references: UNIPROT:P21803; EMBL:X55441; NID:G50141; PIDN:CAA39083.1; PID:G50142
 C;Superfamily: basic fibroblast growth factor receptor 1; immunoglobulin homology; protei
 F;270-343/Domain: immunoglobulin homology <IMM>
 F;478-763/Domain: protein kinase homology <KIN>
 Query Match 79.3%; Score 23; DB 2; Length 820;
 Best Local Similarity 57.1%; Pred. No. 5.8e+02;
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 1 VMXVAEF 7
 Db 638 VMKIADF 644

RESULT 95
 TVHUF2
 fibroblast growth factor receptor 2 precursor - human
 N;Alternate names: fibroblast growth factor receptor K-sam
 N;Contains: protein-tyrosine kinase (EC 2.7.1.112) bek; receptor-like protein-tyrosine k
 C;Species: Homo sapiens (man)
 C;Date: 31-Dec-1993 #sequence revision 19-Jan-1996 #text change 16-Jul-1999
 C;Accession: A42691; B42691; S11691; A36210; B44775; D44775
 R;Kato, M.; Hattori, Y.; Sasaki, H.; Tanaka, M.; Sugano, K.; Yazaki, Y.; Sugimura, T.;
 Proc. Natl. Acad. Sci. U.S.A. 89, 2960-2964, 1992
 A;Title: K-sam gene encodes secreted as well as transmembrane receptor tyrosine kinase.
 A;Reference number: A42691; MUID:92212948; PMID:1313574
 A;Accession: A42691
 A;Molecule type: mRNA
 A;Residues: 1-821 <KAT>
 A;Cross-references: GB:M87770; NID:g186779; PIDN:AAAS9470.1; PID:g186780
 A;Accession: B42691
 A;Molecule type: mRNA
 A;Residues: 1-313,430-821 <KA2>
 A;Cross-references: GB:M87771; NID:g186781; PIDN:AAAS9471.1; PID:g186782
 R;Dionne, C.A.; Crumley, G.; Bellot, F.; Kaplow, J.M.; Searcoss, G.; Ruta, M.; Burgess,
 EMBO J. 9, 2685-2692, 1990
 A;Title: Cloning and expression of two distinct high-affinity receptors cross-reacting w
 A;Reference number: S11691; MUID:90360977; PMID:1697263
 A;Accession: S11691
 A;Molecule type: mRNA
 A;Residues: 1-821 <DIO>
 A;Cross-references: EMBL:X52832; NID:g31373; PIDN:CAA37014.1; PID:g31374
 R;Housaint, E.; Blanquet, P.R.; Champion-Arnaud, P.; Gesnel, M.C.; Torriglia, A.; Court
 Proc. Natl. Acad. Sci. U.S.A. 87, 8180-8184, 1990
 A;Title: Related fibroblast growth factor receptor genes exist in the human genome.
 A;Reference number: A36210; MUID:91045961; PMID:2172978
 A;Accession: A36210
 A;Molecule type: mRNA
 A;Residues: 1-313,'VLK',314-428,431-821 <HOU>
 A;Cross-references: GB:M55614; GB:M37715; NID:g339710; PIDN:AAA61188.1; PID:g339711
 R;Champion-Arnaud, P.; Ronsin, C.; Gilbert, E.; Gesnel, M.C.; Housaint, E.; Breathnach,
 Oncogene 6, 979-987, 1991
 A;Title: Multiple mRNAs code for proteins related to the BEK fibroblast growth factor re
 A;Reference number: A44775; MUID:91296403; PMID:1648704
 A;Accession: B44775
 A;Molecule type: DNA
 A;Residues: 263-361 <CH2>
 A;Cross-references: GB:S40858; NID:g232799; PIDN:AAB19320.1; PID:g232800
 A;Accession: D44775
 A;Molecule type: DNA
 A;Residues: 759-821 <CHA>
 A;Cross-references: GB:S41873
 C;Comment: This receptor binds basic fibroblast growth factor and, with lower affinity,
 C;Genetics:
 A;Gene: GDB:FGFR2
 A;Cross-references: GDB:127273; OMIM:176943
 A;Map position: 10q25.3-10q26
 C;Superfamily: basic fibroblast growth factor receptor 1; immunoglobulin homology; protei
 C;Keywords: alternative splicing; ATP; autophosphorylation; duplication; glycoprotein; g
 protein kinase
 F;1-21/Domain: signal sequence #status predicted <SIG>
 F;22-821/Product: fibroblast growth factor receptor 2 #status predicted <MAT>

F;22-377/Domain: extracellular #status predicted <EXT>
 F;132-138/Region: acidic
 F;172-233/Domain: immunoglobulin homology <IMM>
 F;378-398/Domain: transmembrane #status predicted <TMM>
 F;399-821/Domain: intracellular #status predicted <INT>
 F;479-764/Domain: protein kinase homology <KIN>
 F;487-495/Region: protein kinase ATP-binding motif
 F;62-107,179-231,278-342/bisulfide bonds: #status predicted
 F;83,123,147,241,265,297,318,331/Binding site: carboxydrate (Asn) (covalent) #status pred
 F;517,534,626/Active site: Lys, Glu, Asp #status predicted
 F;631,644/Binding site: magnesium (Asn, Asp) #status predicted
 F;657/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status predicted
 Query Match 79.3%; Score 23; DB 1; Length 821;
 Best Local Similarity 57.1%; Pred. No. 5.8e+02;
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 1 VMXVAEF 7
 Db 639 VMKIADF 645

RESULT 96
 TVMSBK
 fibroblast growth factor receptor bek precursor - mouse
 N;Alternate names: bek transforming protein; fibroblast growth factor receptor 2; keratir
 N;Contains: protein-tyrosine kinase (EC 2.7.1.112) bek
 C;Species: Mus musculus (house mouse)
 C;Date: 30-Jun-1991 #sequence revision 13-Mar-1997 #text change 09-Jul-2004
 C;Accession: A44142; A31378
 R;Mansukhani, A.; Dell'Era, P.; Moscatelli, D.; Kornbluth, S.; Hanafusa, H.; Basilico, C
 Proc. Natl. Acad. Sci. U.S.A. 89, 3305-3309, 1992
 A;Title: Characterization of the murine BEK fibroblast growth factor (FGF) receptor: act
 A;Reference number: A44142; MUID:92228773; PMID:1373495
 A;Accession: A44142
 A;Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tra
 A;Molecule type: mRNA
 A;Residues: 1-821 <MAN>
 A;Cross-references: UNIPROT:P21803; GB:M86441
 R;Kornbluth, S.; Paulson, K.E.; Hanafusa, H.
 Mol. Cell. Biol. 8, 5541-5544, 1988
 A;Title: Novel tyrosine kinase identified by phosphotyrosine antibody screening of cDNA i
 A;Reference number: A31378; MUID:89219016; PMID:2468999
 A;Accession: A31378
 A;Molecule type: mRNA
 A;Residues: 477-821 <KOR>
 A;Cross-references: GB:M23362; NID:G533219; PIDN:AAA37285.1; PID:G533220
 C;Genetics:
 A;Gene: bek
 C;Superfamily: basic fibroblast growth factor receptor 1; immunoglobulin homology; protei
 C;Keywords: ATP; autophosphorylation; duplication; glycoprotein; growth factor receptor;
 F;1-21/Domain: signal sequence #status predicted <SIG>
 F;22-821/Product: fibroblast growth factor receptor bek #status predicted <MAT>
 F;172-233/Domain: immunoglobulin homology <IMM>
 F;378-398/Domain: transmembrane #status predicted <TMM>
 F;479-764/Domain: protein kinase homology <KIN>
 F;487-495/Region: protein kinase ATP-binding motif
 F;62-107,179-231,278-342/bisulfide bonds: #status predicted
 F;83,123,147,241,265,297,318,331/Binding site: carboxydrate (Asn) (covalent) #status pred
 F;517,534,626/Active site: Lys, Glu, Asp #status predicted
 F;657/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status predictor
 Query Match 79.3%; Score 23; DB 1; Length 821;
 Best Local Similarity 57.1%; Pred. No. 5.8e+02;
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 1 VMXVAEF 7
 Db 639 VMKIADF 645

RESULT 97
 TVHUF3

fibroblast growth factor receptor 1 precursor - human
 N;Alternate names: basic fibroblast growth factor receptor; heparin-binding growth factor
 N;Contains: HBGRF 3; HBGRF alpha b1; HBGRF beta a1; HBGRF gamma a1; HBGRF
 C;Species: Homo sapiens (man)
 C;Date: 31-Mar-1989 #sequence_revision 31-Dec-1993 #text change 09-Jul-2004
 C;Accession: S11692; S09226; A61536; S26739; S26738; S17374; S17375; S17
 R;Dionne, C.A.; Crumley, G.; Bellot, F.; Kaplow, J.M.; Searfoss, G.; Ruta, M.; Burgess,
 EMBO J. 9, 2695-2692, 1990
 A;Title: Cloning and expression of two distinct high-affinity receptors cross-reacting w
 A;Reference number: S11691; MUID:90360977; PMID:1697263
 A;Accession: S11692
 A;Molecule type: mRNA
 A;Residues: 1-822 <DIO>
 A;Cross-references: UNIPROT:P11362; EMBL:X52833; NID:g31377; PIDN:CAA37015.1; PID:g31378
 R;Hou, J.; Kan, M.; McKeenhan, K.; McBride, G.; Adams, P.; McKeenhan, W.L.
 Science 251, 665-668, 1991
 A;Title: Fibroblast growth factor receptors from liver vary in three structural domains.
 A;Reference number: A40862; MUID:91126480; PMID:1846977
 A;Accession: B40862
 A;Molecule type: mRNA
 A;Residues: 1-822 <HOU>
 A;Cross-references: GB:M63887
 A;Note: potentially twelve variants may arise by alternative splicing in liver cells
 R;Tsacchi, A.; Bergonzoni, L.; Sarmientos, P.
 Nucleic Acids Res. 18, 1906, 1990
 A;Title: Complete sequence of a human receptor for acidic and basic fibroblast growth fa
 A;Reference number: S09226; MUID:90245600; PMID:2159626
 A;Accession: S09226
 A;Molecule type: mRNA
 A;Residues: 1-816, 'R', 818-822 <ISA>
 A;Cross-references: EMBL:X51803; NID:g31367; PIDN:CAA36101.1; PID:g31368
 R;Wannstrom, S.; Sandstrom, C.; Claesson-Welsh, L.
 Growth Factors 4, 197-208, 1991
 A;Title: cDNA cloning and expression of a human FGF receptor which binds acidic and basi
 A;Reference number: A61533; MUID:92118394; PMID:1722683
 A;Accession: A61533
 A;Molecule type: mRNA
 A;Residues: 1-147, 150-193, 'S', 195-822 <WEN>
 A;Cross-references: GB:M34641; NID:g182529; PIDN:AAA5835.1; PID:g182530
 A;Experimental source: teracarcinoma cell line Tera-2
 R;Kiefer, M.C.; Baird, A.; Nguyen, T.; George-Nascimento, C.G.; Mason, O.B.; Boley, L.J.
 Growth Factors 5, 115-127, 1991
 A;Title: Molecular cloning of a human basic fibroblast growth factor receptor cDNA and e
 A;Reference number: A61536; MUID:92118399; PMID:1662973
 A;Accession: A61536
 A;Status: not compared with conceptual translation
 A;Molecule type: mRNA
 A;Residues: 1-147, 150-822 <KIE>
 R;Tronick, S.R.
 submitted to the EMBL Data Library, January 1991
 A;Reference number: S19167
 A;Accession: S26739
 A;Status: preliminary
 A;Molecule type: mRNA
 A;Residues: 1-822 <TR1>
 A;Cross-references: EMBL:X57121; NID:g31392; PIDN:CAA40403.1; PID:g31393
 A;Accession: S26738
 A;Molecule type: mRNA
 A;Residues: 1-147, 150-822 <TR2>
 A;Cross-references: EMBL:X57120; NID:g31390; PIDN:CAA40402.1; PID:g31391
 R;Risemann, A.; Ahn, J.A.; Graziani, G.; Tronick, S.R.; Ron, D.
 Oncogene 6, 1195-1202, 1991
 A;Title: Alternative splicing generates at least five different isoforms of the human ba
 A;Reference number: S17373; MUID:91319400; PMID:1650441
 A;Accession: S17374
 A;Molecule type: mRNA
 A;Residues: 1-30, 120-147, 150-822 <E14>
 A;Cross-references: EMBL:X57122; NID:g31386; PIDN:CAA40404.1; PID:g31387
 A;Note: the sequence from Fig. 3 is inconsistent with that from Fig. 1 in lacking 25-Pro
 A;Note: this form is designated isoform I
 A;Accession: S17375
 A;Molecule type: mRNA
 A;Residues: 1-30, 120-822 <E13>

A;Cross-references: EMBL:X57119; NID:g31388; PIDN:CAA40401.1; PID:g31389
 A;Note: the sequence from Fig. 3 is inconsistent with that from Fig. 1 in lacking 25-Pro;
 A;Note: this form is designated isoform II
 A;Accession: S17377
 A;Molecule type: mRNA
 A;Residues: 1-90, 'D', '91-822 <E12>
 A;Cross-references: EMBL:X57121
 A;Note: the sequence from Fig. 3 is inconsistent with that from Fig. 1 in lacking 25-Pro;
 A;Note: this form is designated isoform III
 A;Accession: S17376
 A;Molecule type: mRNA
 A;Residues: 1-90, 'D', '91-147, 150-822 <E15>
 A;Cross-references: EMBL:X57120
 A;Note: this form is designated isoform IV
 R;Itou, N.; Terachi, T.; Ohta, M.; Seo, M.K.
 Biochem. Biophys. Res. Commun. 169, 680-685, 1990
 A;Title: The complete amino acid sequence of the shorter form of human basic fibroblast g
 A;Reference number: A35479; MUID:90290512; PMID:2162671
 A;Accession: A35479
 A;Molecule type: mRNA
 A;Residues: 1-30, 120-147, 150-468, 'L', 470-822 <ITO>
 A;Cross-references: GB:M37722; NID:g179413; PIDN:AAA75007.1; PID:g179415
 A;Note: both the longer and shorter forms are expressed in the placenta
 R;Johnson, D.E.; Lee, P.L.; Lu, J.; Williams, L.T.
 Mol. Cell. Biol. 10, 4728-4736, 1990
 A;Title: Diverse forms of a receptor for acidic and basic fibroblast growth factors.
 A;Reference number: A36464; MUID:90355989; PMID:2167437
 A;Accession: A36464
 A;Molecule type: mRNA
 A;Residues: 1-30, 120-822 <JOH>
 A;Cross-references: GB:M34185; NID:g182531; PIDN:AAA5836.1; PID:g182532
 A;Accession: B36464
 A;Molecule type: mRNA
 A;Residues: 1-30, 120-147, 150-191, 'E', 193-822 <JO2>
 A;Cross-references: GB:M34186; NID:g182533; PIDN:AAA5837.1; PID:g182534
 A;Accession: E36464
 A;Status: not compared with conceptual translation
 A;Molecule type: DNA
 R;Ruta, M.; Howk, R.; Ricca, G.; Drohan, W.; Zabelshansky, M.; Laureys, G.; Barton, D.B.;
 Oncogene 3, 9-15, 1988
 A;Title: A novel protein tyrosine kinase gene whose expression is modulated during endot
 A;Reference number: A28361
 A;Accession: A28361
 A;Molecule type: mRNA
 A;Residues: 201-671, 'IYLTGS', 677-822 <RUT>
 A;Cross-references: EMBL:Y00665
 R;Hattori, Y.; Odagiri, H.; Katoh, O.; Sakamoto, H.; Morita, T.; Shimotohno, K.; Tobinai,
 Cancer Res. 52, 3367-3371, 1992
 A;Title: K-sam-related gene, N-sam, encodes fibroblast growth factor receptor and is exp
 A;Reference number: S25420; MUID:92282615; PMID:1317750
 A;Accession: S25420
 A;Status: preliminary
 A;Molecule type: mRNA
 A;Residues: 1-822 <HAT>
 A;Cross-references: EMBL:X66945; NID:g35109; PIDN:CAA47375.1; PID:g35110
 R;Rushnati, M.; Coltrini, D.; Caccia, P.; Bell'Era, P.; Zoppetti, G.; Oreste, P.; Valsasir
 Biochem. Biophys. Res. Commun. 203, 450-458, 1994
 A;Title: Distinct role of 2-O-, N-, and 6-O-sulfate groups of heparin in the formation of
 A;Reference number: PC2394; MUID:94354840; PMID:8074689
 A;Accession: PC2394
 A;Molecule type: protein
 A;Residues: 81-100 <RUS>
 A;Experimental source: recombinant soluble form of extracellular domain after expression
 A;Note: this sequence represents the amino end of a 33k fragment protected from trypsin
 C;Comment: This receptor binds acidic and basic fibroblast growth factors with high affir
 C;Genetics:
 A;Gene: GDB:FGFR1; FLT2
 A;Cross-references: GDB:119913; OMIM:136350
 A;Map position: 8p11.2-8p11.1
 C;Superfamily: basic fibroblast growth factor receptor 1; immunoglobulin homology; protei
 C;Keywords: alternative splicing; ATP; autophosphorylation; duplication; glycoprotein; g
 tyrosine-specific protein kinase

F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-822/Product: fibroblast growth factor receptor 1, form alpha a1 #status predicted <F>
F;22-427,430-822/Product: fibroblast growth factor receptor 1, form alpha b1 #status predicted <F>
F;22-376/Domain: extracellular #status predicted <EXT>
F;22-147,150-822/Product: fibroblast growth factor receptor 1, form 3 #status predicted <F>
F;22-30,120-822/Product: fibroblast growth factor receptor 1, form beta a1 #status predicted <F>
F;22-30,120-822/Product: fibroblast growth factor receptor 1, form beta b1 #status predicted <F>
F;22-30,120-427,430-822/Product: fibroblast growth factor receptor 1, form beta b1 #status predicted <F>
F;126-133/Region: acidic
F;161-822/Product: fibroblast growth factor receptor 1, form gamma a1 #status predicted <F>
F;161-427,430-822/Product: fibroblast growth factor receptor 1, form gamma b1 #status predicted <F>
F;171-232/Domain: immunoglobulin homology <IMM>
F;377-397/Domain: transmembrane #status predicted <TM>
F;398-822/Domain: intracellular #status predicted <INT>
F;476-761/Domain: protein kinase homology <KIN>
F;484-492/Region: protein kinase ATP-binding motif
F;55-101,178-230,277-341/Disulfide bonds: #status predicted <F>
F;77,117,227,240,264,296,317,330/Binding site: carboxylate (Asn) (covalent) #status predicted <F>
F;514,531,623/Active site: Lys, Glu, Asp #status predicted <F>
F;628,641/Binding site: magnesium (Asn, Asp) #status predicted <F>
F;654/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status predicted <F>

Query Match 79.3%; Score 23; DB 1; Length 822;
Best Local Similarity 57.1%; Pred. No. 5.8e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
||| :||
Db 636 VMKIADF 642

RESULT 98
TVMSFG
N;Alternate names: basic fibroblast growth factor precursor - mouse
N;Contains: protein-tyrosine kinase (EC 2.7.1.112) flg
C;Species: Mus musculus (house mouse)
C;Date: 31-Dec-1993 #sequence revision 31-Dec-1993 #text change 09-Jul-2004
C;Accession: A34849; B34849; S09953; A35794; A43025; PC2277
R;Reid, H.H.; Wilks, A.F.; Bernard, O.
Proc. Natl. Acad. Sci. U.S.A. 87, 1596-1600, 1990
A;Title: Two forms of the basic fibroblast growth factor receptor-like mRNA are expressed in rat 3Y1 fibroblasts
A;Reference number: A34849; MUID:90160373; PMID:1689490
A;Accession: A34849
A;Molecule type: mRNA
A;Residues: 1-822 <RE1>
A;Cross-references: UNIPROT:P16092; GB:M28998; NID:G192164; PIDN:AAA37290.1; PID:G309116
A;Accession: B34849
A;Molecule type: mRNA
A;Residues: 1-30,120-822 <RE2>
A;Cross-references: GB:M28998
A;Note: the shorter form is expressed at higher levels in neuronal cells at early stages of development
R;Safran, A.; Avivi, A.; Orr-Urtreger, A.; Neufeld, G.; Lonai, P.; Givol, D.; Yarden, Y.
Oncogene 5, 635-643, 1990
A;Title: The murine flg gene encodes a receptor for fibroblast growth factor.
A;Reference number: S09953; MUID:90265603; PMID:2161096
A;Accession: S09953
A;Molecule type: mRNA
A;Residues: 1-147,150-255, 'ILQ', 259-439, 'A', 441-755, 'R', 757-822 <SAF>
A;Cross-references: EMBL:X51893; NID:950959; PIDN:CAA36175.1; PID:G50360
R;Mansukhani, A.; Moscatelli, D.; Talarico, D.; Levytska, V.; Basilio, C.
Proc. Natl. Acad. Sci. U.S.A. 87, 4378-4382, 1990
A;Title: A murine fibroblast growth factor (FGF) receptor expressed in CHO cells is active in the presence of a specific ligand
A;Reference number: A35794; MUID:90272715; PMID:2161540
A;Accession: A35794
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-30,120-228, 'S', 230-255, 'ILQ', 259-269, 'A', 271-543, 'M', 545-628, 'LV', 631-755, 'R', 757-822 <SAR>
A;Cross-references: GB:M33760
R;Basilico, C.
submitted to GenBank, March 1990
A;Reference number: A43025
A;Accession: A43025
A;Molecule type: mRNA

A;Residues: 1-30,120-228, 'S', 230-255, 'ILQ', 259-269, 'A', 271-543, 'M', 545-755, 'R', 757-764, 'I', 765-822 <SAR>
A;Cross-references: GB:M33760; NID:G193298; PIDN:AAA37622.1; PID:G309240
R;Harada, T.; Saito, H.; Kouhara, H.; Kurebayashi, S.; Kasayama, S.; Terakawa, N.; Kishimoto, A.
Biochem. Biophys. Res. Commun. 205, 1057-1063, 1994
A;Title: Murine fibroblast growth factor receptor 1 gene generates multiple messenger RNAs
A;Reference number: PC2277; MUID:95100926; PMID:7802632
A;Accession: PC2277
A;Molecule type: DNA
A;Residues: 1-15 <HAR>
A;Cross-references: GB:S74765; NID:G833887; PIDN:AAB32845.1; PID:G833889
C;Comment: This protein mediates the biological actions of heparin-binding growth factor receptor 1
C;Superfamily: basic fibroblast growth factor receptor 1; immunoglobulin homology; protein kinase
C;Keywords: alternative splicing; ATP; autophosphorylation; duplication; glycoprotein; growth factor receptor
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-822/Product: fibroblast growth factor receptor 1 #status predicted <MAT>
F;22-376/Domain: extracellular #status predicted <EXT>
F;22-30,120-822/Product: fibroblast growth factor receptor 1, short form #status predicted <F>
F;126-133/Region: acidic
F;171-232/Domain: immunoglobulin homology <IMM>
F;377-397/Domain: transmembrane #status predicted <TM>
F;398-822/Domain: intracellular #status predicted <INT>
F;476-761/Domain: protein kinase homology <KIN>
F;484-492/Region: protein kinase ATP-binding motif
F;55-101,178-230,277-341/Disulfide bonds: #status predicted <F>
F;77,117,227,240,264,296,317,330/Binding site: carboxylate (Asn) (covalent) #status predicted <F>
F;514,531,623/Active site: Lys, Glu, Asp #status predicted <F>
F;628,641/Binding site: magnesium (Asn, Asp) #status predicted <F>
F;654/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status predicted <F>

Query Match 79.3%; Score 23; DB 1; Length 822;
Best Local Similarity 57.1%; Pred. No. 5.8e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
||| :||
Db 636 VMKIADF 642

RESULT 99
B54846
fibroblast growth factor receptor b precursor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 21-Jul-1995 #sequence revision 28-Jul-1995 #text change 16-Jul-1999
C;Accession: B54846
R;Takagi, Y.; Shrivastav, S.; Miki, T.; Sakaguchi, K.
J. Biol. Chem. 269, 23743-23749, 1994
A;Title: Molecular cloning and expression of the acidic fibroblast growth factor receptor 1
A;Reference number: A54846; MUID:94375484; PMID:8089146
A;Accession: B54846
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-822 <RA>
C;Superfamily: basic fibroblast growth factor receptor 1; immunoglobulin homology; protein kinase
C;Keywords: ATP; growth factor receptor
F;172-233/Domain: immunoglobulin homology <IMM>
F;480-765/Domain: protein kinase homology <KIN>
F;488-496/Region: protein kinase ATP-binding motif

Query Match 79.3%; Score 23; DB 2; Length 822;
Best Local Similarity 57.1%; Pred. No. 5.8e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
||| :||
Db 640 VMKIADF 646

RESULT 100
A45081
fibroblast growth factor receptor 2 IIIB, FGFR2 IIIB - human
N;Contains: protein-tyrosine kinase (EC 2.7.1.112)

C;Species: Homo sapiens (man)
C;Date: 17-Feb-1994 #sequence_revision 17-feb-1994 #text_change 09-Jul-2004
C;Accession: A45081
R;Bell, K.R.; Williams, L.T.
J. Biol. Chem. 267, 21225-21229, 1992
A;Title: A novel form of fibroblast growth factor receptor 2. Alternative splicing of the
A;Reference number: A45081; MUID:93016048; PMID:1400433
A;Accession: A45081
A;Status: preliminary
A;Molecule type: nucleic acid
A;Residues: 1-822
A;Cross-references: UNIPROT:P21802; GB:M97193; NID:g182566; PIDN:AAA52449.1; PID:g182567
A;Experimental source: placenta
A;Note: sequence extracted from NCBI backbone (NCBI:P116207)
C;Superfamily: basic fibroblast growth factor receptor 1; immunoglobulin homology; prote
C;Keywords: ATP; autophosphorylation; duplication; glycoprotein; growth factor receptor;
F;172-233/Domain: immunoglobulin homology <IMM>
F;379-399/Domain: transmembrane #status predicted <TMM>
F;480-765/Domain: protein kinase homology <KIN>
F;488-496/Region: protein kinase ATP-binding motif
F;62-107,179-231,278-340/Disulfide bonds: #status predicted
F;518,535,627/Active site: Lys, Glu, Asp #status predicted

Query Match 79.3%; Score 23; DB 2; Length 822;
Best Local Similarity 57.1%; Pred. No. 5.8e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VMXVAEF 7
|| :||
Db 640 VMKIADF 646

Search completed: June 13, 2005, 14:01:46
Job time : 28 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2005, 13:23:36 ; Search time 113 Seconds
(without alignments)
31.722 Million cell updates/sec

Title: 09730329-60ED
Perfect score: 29
Sequence: 1 VMXVAF 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 300 summaries

Database :

UniProt_03:*

1: uniprot_sprot:*

2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	28	96.6	776	2 Q94IL7	Q94IL7 populus tre
2	28	96.6	3414	1 POLG LANVT	P29837 l genome po
3	28	96.6	3414	2 Q9IG39	Q9IG39 langat viru
4	28	96.6	3414	2 Q9IG40	Q9IG40 langat viru
5	27	93.1	433	2 Q7XA44	Q7XA44 solanum bul
6	27	93.1	775	1 POT3 ARATH	POT3 arabisidopsi
7	27	93.1	814	2 Q8VXQ3	Q8VXQ3 cymodocea n
8	26	89.7	223	1 DEOC MYCGE	P47296 mycoplasma
9	26	89.7	224	1 DEOC MYCPN	P09924 mycoplasma
10	26	89.7	262	2 Q8U0A9	Q8U0A9 pyrococcus
11	26	89.7	324	2 Q92PQ5	Q92PQ5 arabidopsis
12	26	89.7	479	2 Q65892	Q65892 bacillus li
13	26	89.7	731	2 Q6Z5S4	Q6Z5S4 oryza sativ
14	26	89.7	821	2 Q9XBW4	Q9XBW4 porphyromon
15	25	86.2	205	2 Q62IK2	Q62IK2 burkholderi
16	25	86.2	205	2 Q63VR6	Q63VR6 burkholderi
17	25	86.2	253	1 YTBQ_BACSU	P53560 bacillus su
18	25	86.2	284	2 Q8U3Z7	Q8U3Z7 pyrococcus
19	25	86.2	406	2 Q82Z77	Q82Z77 enterococcu
20	25	86.2	617	2 Q7VMA4	Q7VMA4 haemophilus
21	25	86.2	632	1 ETED SCHPO	P87111 s probable
22	25	86.2	633	2 Q7UY37	Q7UY37 rhodopirell
23	25	86.2	1077	2 P92974	P92974 arabidopsis
24	24	82.8	140	2 Q6Y252	Q6Y252 pagrus majo
25	24	82.8	145	1 RL13 HALMA	P29198 halocaula
26	24	82.8	148	2 Q72IL7	Q72IL7 thermus the
27	24	82.8	149	2 Q8TLX8	Q8TLX8 methanosarc
28	24	82.8	154	2 Q86ZD3	Q86ZD3 fusarium sp
29	24	82.8	157	1 RAP_TAROF	Q49065 taraxacum o
30	24	82.8	161	2 Q89G21	Q89G21 bradyrhizob
31	24	82.8	162	2 Q6AQ12	Q6AQ12 desulfotale

32	24	82.8	165	2 Q8SRD1	Q8SRD1 encephalito
33	24	82.8	169	1 SSB_RHOBA	P59932 rhodopirell
34	24	82.8	169	2 Q6M2B7	Q6M2B7 corynebacte
35	24	82.8	189	2 Q9N4I6	Q9N4I6 caenorhabdi
36	24	82.8	189	2 Q97F65	Q97F65 clostridium
37	24	82.8	206	2 Q8XWK7	Q8XWK7 ralsstridia s
38	24	82.8	236	1 YP58_MYCTU	P65005 mycobacteri
39	24	82.8	236	1 YP58_MYCBO	P65006 mycobacteri
40	24	82.8	238	2 Q6NIE1	Q6NIE1 corynebacte
41	24	82.8	248	2 Q97008	Q97008 leishmania
42	24	82.8	249	2 Q9PKD5	Q9PKD5 chlamydia m
43	24	82.8	270	2 Q9NE74	Q9NE74 leishmania
44	24	82.8	297	2 Q661U0	Q661U0 borrelia ga
45	24	82.8	317	2 Q9U375	Q9U375 caenorhabdi
46	24	82.8	335	1 K6PF_STRPN	Q97RC6 streptococc
47	24	82.8	335	1 K6PF_STRR6	Q8DQ5 streptococc
48	24	82.8	336	2 Q88SF9	Q88SF9 lactobacill
49	24	82.8	337	2 Q63TY5	Q63TY5 burkholderi
50	24	82.8	367	2 Q97RK0	Q97RK0 clostridium
51	24	82.8	375	2 Q7Q854	Q7Q854 anopheles g
52	24	82.8	378	2 Q8L789	Q8L789 arabidopsis
53	24	82.8	378	2 Q9FNG6	Q9FNG6 arabidopsis
54	24	82.8	414	2 Q6K7P6	Q6K7P6 oryza sativ
55	24	82.8	419	2 Q67UR2	Q67UR2 oryza sativ
56	24	82.8	419	2 Q7M8F1	Q7M8F1 wolinnella s
57	24	82.8	422	2 Q8KI64	Q8KI64 pseudomonas
58	24	82.8	425	2 Q7P4E5	Q7P4E5 fusobacteri
59	24	82.8	426	2 Q6UD01	Q6UD01 uncultured
60	24	82.8	426	2 Q6UD29	Q6UD29 uncultured
61	24	82.8	439	2 Q8RG30	Q8RG30 fusobacteri
62	24	82.8	442	2 Q7XA19	Q7XA19 solanum bul
63	24	82.8	457	2 Q6ZS89	Q6ZS89 homo sapien
64	24	82.8	463	2 Q8WZNS	Q8WZNS pleurotus o
65	24	82.8	469	1 G64F_DROME	P83297 drosophila
66	24	82.8	484	2 Q9C7R2	Q9C7R2 arabidopsis
67	24	82.8	565	1 TP6B_AERPE	Q5YE4 aeropyrum p
68	24	82.8	582	2 Q58843	Q58843 pyrococcus
69	24	82.8	584	2 Q9UZNO	Q9UZNO pyrococcus
70	24	82.8	619	2 Q8SRA9	Q8SRA9 encephalito
71	24	82.8	659	2 Q8P8F1	Q8P8F1 xanthomonas
72	24	82.8	738	2 P91063	P91063 caenorhabdi
73	24	82.8	871	2 Q846V6	Q846V6 staphylococ
74	24	82.8	916	1 SYI_STAEP	Q8CSX1 staphylococ
75	24	82.8	921	1 SYI_BACSU	Q45477 bacillus su
76	24	82.8	921	2 Q63ED1	Q63ED1 bacillus ce
77	24	82.8	921	2 Q732H4	Q732H4 bacillus ce
78	24	82.8	921	2 Q819R4	Q819R4 bacillus ce
79	24	82.8	921	2 Q81WE4	Q81WE4 bacillus an
80	24	82.8	921	2 Q9K9V0	Q9K9V0 bacillus ha
81	24	82.8	921	2 Q6HER9	Q6HER9 bacillus th
82	24	82.8	922	2 Q65JV7	Q65JV7 bacillus li
83	24	82.8	961	2 Q6G4C0	Q6G4C0 bartonella
84	24	82.8	971	1 AMPN_HAECO	Q10737 haemochus
85	24	82.8	972	2 Q86G73	Q86G73 haemochus
86	24	82.8	1095	1 AT9B_HUMAN	Q43861 homo sapien
87	24	82.8	1095	1 AT9B_MOUSE	P98195 mus musculu
88	24	82.8	1122	1 VAB1_CAEEL	Q61460 caenorhabdi
89	24	82.8	1130	2 Q965S4	Q965S4 caenorhabdi
90	24	82.8	1146	2 Q68FM3	Q68FM3 mus musculu
91	24	82.8	1173	2 Q7NEB4	Q7NEB4 anopheles g
92	24	82.8	1566	2 Q7NEB4	Q7NEB4 gloeobacter
93	24	82.8	2033	2 Q7X195	Q7X195 oryza sativ
94	24	82.8	2201	2 Q70KH4	Q70KH4 streptomyc
95	23	79.3	53	2 Q712V2	Q712V2 drosophila
96	23	79.3	56	2 Q91995	Q91995 xenopus lae
97	23	79.3	56	2 Q788X1	Q788X1 xenopus lae
98	23	79.3	57	2 Q9QW56	Q9QW56 mus sp. p
99	23	79.3	57	2 Q08556	Q08556 coturnix co
100	23	79.3	61	2 Q9P166	Q9P166 homo sapien
101	23	79.3	68	2 Q9Y121	Q9Y121 gallus gall
102	23	79.3	89	2 Q74YJ4	Q74YJ4 yersinia pe
103	23	79.3	89	2 Q9ZGY2	Q9ZGY2 yersinia pe
104	23	79.3	96	2 Q8U8Q1	Q8U8Q1 agrobacteri

105	23	79.3	102	2	Q741L6	Q741L6 mycobacteri	178	23	79.3	346	2	Q70MC3	Q70mc3 staphylococ
106	23	79.3	105	2	Q9L6S7	Q9L6S7 salmonella	179	23	79.3	348	2	Q8CBY7	Q8cbY7 mus musculus
107	23	79.3	112	2	Q8X7L7	Q8X7L7 escherichia	180	23	79.3	350	2	Q8IGU7	Q8igu7 drosophila
108	23	79.3	119	2	Q8S9E0	Q8S9E0 pseudomonas	181	23	79.3	350	2	Q9NGZ2	Q9ngz2 drosophila
109	23	79.3	123	2	Q8TW74	Q8TW74 methanopyru	182	23	79.3	350	2	Q9V577	Q9v577 drosophila
110	23	79.3	126	2	Q7WFX5	Q7WFX5 bordetella	183	23	79.3	352	1	CYCR_ROSDE	P26278 rosebacter
111	23	79.3	127	2	Q73KS9	Q73KS9 treponema d	184	23	79.3	352	2	Q83053	Q83053 louping ill
112	23	79.3	129	2	Q63000	Q63000 rattus norv	185	23	79.3	354	2	Q64L14	Q64l14 nemertoderm
113	23	79.3	131	2	Q8WMP3	Q8wmp3 equus cabal	186	23	79.3	362	2	Q9ULZ2	Q9ulz2 caenorhabdi
114	23	79.3	136	2	Q6L153	Q6L153 picophilus	187	23	79.3	362	2	Q8PEB8	Q8peb8 xanthomonas
115	23	79.3	138	2	Q89XN0	Q89XN0 bradyrhizob	188	23	79.3	370	2	Q886Z4	Q886z4 rhizobium l
116	23	79.3	138	2	Q9PSF8	Q9psf8 gallus gall	189	23	79.3	373	2	Q66AS2	Q66as2 versinia ps
117	23	79.3	143	1	RISB_ARCFU	Q28152 archaeoglob	190	23	79.3	373	2	Q8ZES8	Q8zes8 versinia pe
118	23	79.3	143	2	Q7PQL4	Q7pql4 anopheles g	191	23	79.3	374	2	Q63240	Q63240 rattus norv
119	23	79.3	145	2	Q9BHM6	Q9bhm6 rhizobium l	192	23	79.3	377	2	Q88482	Q88482 tick-borne
120	23	79.3	154	2	Q9PHY3	Q9phy3 arabidopsis	193	23	79.3	381	2	Q6RCD0	Q6rcd0 pseudomonas
121	23	79.3	157	2	Q77807	Q77807 bos taurus	194	23	79.3	385	2	Q6M6P3	Q6m6p3 corynebacte
122	23	79.3	157	2	Q28111	Q28111 bos taurus	195	23	79.3	386	2	Q8D0F7	Q8d0f7 versinia pe
123	23	79.3	157	2	Q28332	Q28332 callithrix	196	23	79.3	387	2	Q9HYM7	Q9hyM7 pseudomonas
124	23	79.3	157	2	Q9NOK5	Q9nok5 sus scrofa	197	23	79.3	388	2	Q91146	Q91146 notophthalm
125	23	79.3	157	2	Q69K45	Q69k45 oryza sativ	198	23	79.3	388	2	Q8KH27	Q8kh27 helicobacte
126	23	79.3	162	1	AROK_LACLA	Q9ceul lactococcus	199	23	79.3	389	2	Q8KH91	Q8kh91 helicobacte
127	23	79.3	165	2	Q9F4D2	Q9f4d2 bacteroides	200	23	79.3	389	2	Q8KHB8	Q8kHb8 helicobacte
128	23	79.3	168	2	Q80YG0	Q80yg0 cavia porce	201	23	79.3	389	2	Q8LON2	Q8lOn2 helicobacte
129	23	79.3	169	2	Q8URF5	Q8urf5 agrobacteri	202	23	79.3	389	2	Q8LON3	Q8lOn3 helicobacte
130	23	79.3	172	2	Q02528	Q02528 oryzias lat	203	23	79.3	389	2	Q8LON4	Q8lOn4 helicobacte
131	23	79.3	172	2	Q02529	Q02529 oryzias lat	204	23	79.3	389	2	Q8LON5	Q8lOn5 helicobacte
132	23	79.3	172	2	Q91176	Q91176 oryzias lat	205	23	79.3	389	2	Q8LON6	Q8lOn6 helicobacte
133	23	79.3	172	2	Q91177	Q91177 oryzias lat	206	23	79.3	389	2	Q8LON7	Q8lOn7 helicobacte
134	23	79.3	173	2	Q69RN2	Q69rn2 oryza sativ	207	23	79.3	389	2	Q8LON8	Q8lOn8 helicobacte
135	23	79.3	184	2	Q675T9	Q675t9 oikopleura	208	23	79.3	389	2	Q8LON9	Q8lOn9 helicobacte
136	23	79.3	185	2	Q69T19	Q69t19 oryza sativ	209	23	79.3	389	2	Q8LOP0	Q8lOp0 helicobacte
137	23	79.3	189	2	Q8FC12	Q8fc12 escherichia	210	23	79.3	389	2	Q8LOP1	Q8lOp1 helicobacte
138	23	79.3	191	2	Q9KKV1	Q9kkv1 vibrio chol	211	23	79.3	389	2	Q8LOP2	Q8lOp2 helicobacte
139	23	79.3	192	2	Q8L647	Q8l647 deschampsia	212	23	79.3	389	2	Q8LOP3	Q8lOp3 helicobacte
140	23	79.3	203	2	Q898Y8	Q898y8 clostridium	213	23	79.3	389	2	Q8LOP4	Q8lOp4 helicobacte
141	23	79.3	204	2	Q7Q4G0	Q7q4g0 anopheles g	214	23	79.3	389	2	Q8LOP5	Q8lOp5 helicobacte
142	23	79.3	205	2	Q8TY17	Q8ty17 methanopyru	215	23	79.3	389	2	Q8LOP6	Q8lOp6 helicobacte
143	23	79.3	206	2	Q6BMY0	Q6bmy0 debaryomyce	216	23	79.3	389	2	Q8LOP7	Q8lOp7 helicobacte
144	23	79.3	206	2	Q6Z056	Q6z056 oryza sativ	217	23	79.3	389	2	Q8LOP8	Q8lOp8 helicobacte
145	23	79.3	209	2	Q7C215	Q7c215 agrobacteri	218	23	79.3	393	2	Q9V088	Q9v088 helicobacte
146	23	79.3	209	2	Q89L29	Q89l29 bradyrhizob	219	23	79.3	394	2	Q9PVV3	Q9pvv3 lampetra re
147	23	79.3	221	2	Q8DU18	Q8du18 streptococc	220	23	79.3	395	2	Q70W10	Q70w10 ciona intes
148	23	79.3	230	2	Q8NS24	Q8ns24 corynebacte	221	23	79.3	395	2	Q8VPP2	Q8vpp2 micrococcus
149	23	79.3	244	2	Q7PMW1	Q7pmw1 anopheles g	222	23	79.3	398	2	Q8RQH7	Q8rqh7 listeria mo
150	23	79.3	248	2	Q8FR32	Q8fr32 corynebacte	223	23	79.3	399	2	Q8Y674	Q8y674 listeria mo
151	23	79.3	248	2	Q9PIU8	Q9piu8 campylobact	224	23	79.3	399	2	Q92AI3	Q92ai3 listeria in
152	23	79.3	249	1	AQP_AEDAE	Q9nhw7 aedes aegyp	225	23	79.3	399	2	Q71VJ1	Q71vj1 listeria mo
153	23	79.3	254	1	PMW_YEAST	P07283 saccharomyc	226	23	79.3	402	2	Q6MLA8	Q6mla8 bdellovibri
154	23	79.3	254	2	Q70D76	Q70d76 saccharomyc	227	23	79.3	409	1	HDRD_METMA	Q8pww3 methanosarc
155	23	79.3	254	2	Q70D77	Q70d77 saccharomyc	228	23	79.3	412	2	Q95Z02	Q95z02 toxoplasma
156	23	79.3	254	2	Q88A64	Q88a64 pseudomonas	229	23	79.3	413	2	Q65LK6	Q65lk6 bacillus li
157	23	79.3	275	1	NULM_MYTED	Q00860 mytilus edu	230	23	79.3	414	2	Q62J32	Q62jj2 burkholderi
158	23	79.3	276	2	Q88A71	Q88a71 pseudomonas	231	23	79.3	415	2	Q64334	Q64334 rattus norv
159	23	79.3	280	2	Q8NRY7	Q8nry7 corynebacte	232	23	79.3	418	2	Q63T78	Q63t78 burkholderi
160	23	79.3	291	1	AMPM_ARCFU	Q28438 archaeoglob	233	23	79.3	421	1	MUAI_STAAM	Q93ln5 staphylococ
161	23	79.3	299	1	RL22_DROME	P50887 drosophila	234	23	79.3	421	1	MUAI_STAAN	P84U58 staphylococ
162	23	79.3	302	2	Q645Z3	Q645z3 gorilla gor	235	23	79.3	421	1	MUAI_STAAP	P84U59 staphylococ
163	23	79.3	305	2	Q68SR5	Q68sr5 mytilus edu	236	23	79.3	421	1	MUAI_STAEP	Q8crn6 staphylococ
164	23	79.3	310	2	Q7TRJ9	Q7trj9 mus musculu	237	23	79.3	421	2	Q6G7L0	Q6g7l0 staphylococ
165	23	79.3	312	2	Q6LX35	Q6lx35 methanococc	238	23	79.3	421	2	Q6GEX5	Q6gex5 staphylococ
166	23	79.3	312	2	Q9UANI	Q9uani drosophila	239	23	79.3	439	2	Q7UG05	Q7ug05 rhodospirell
167	23	79.3	312	2	Q923Q8	Q923q8 mus musculu	240	23	79.3	440	2	Q9F5V2	Q9f5v2 bacteroides
168	23	79.3	314	2	Q7RTR8	Q7rtr8 homo sapien	241	23	79.3	441	2	Q638K1	Q638k1 bacillus ce
169	23	79.3	314	2	Q645X0	Q645x0 homo sapien	242	23	79.3	441	2	Q734J7	Q734j7 bacillus ce
170	23	79.3	314	2	Q8C053	Q8c053 staphylococ	243	23	79.3	441	2	Q81B22	Q81b22 bacillus ce
171	23	79.3	321	2	Q7X1G1	Q7x1g1 leptospiril	244	23	79.3	441	2	Q81M20	Q81m20 bacillus an
172	23	79.3	324	2	Q8HQH2	Q8hqj2 ornithodor	245	23	79.3	441	2	Q6HG41	Q6hg41 bacillus th
173	23	79.3	325	2	Q7X1N1	Q7x1n1 leptospiril	246	23	79.3	442	2	Q19704	Q19704 caenorhabdi
174	23	79.3	329	2	Q9U8W3	Q9u8w3 brachyost	247	23	79.3	453	2	Q23615	Q23615 caenorhabdi
175	23	79.3	333	2	Q89EK5	Q89ek5 bradyrhizob	248	23	79.3	467	2	Q97G30	Q97g30 clostridium
176	23	79.3	333	2	Q9U8W4	Q9u8w4 eptatretus	249	23	79.3	479	2	Q7Tsf9	Q7tsf9 mus musculu
177	23	79.3	335	2	Q8TX93	Q8tx93 methanopyru	250	23	79.3	487	2	Q8G1Z9	Q8g1z9 brucella su

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR ENBL; M73835; AAA02740.1; ALT_TERM.
 DR ENBL; S35365; AAB22165.1; -.
 DR PIR; A42545; A42545.
 DR HSSP; P14336; 1SVB.
 DR MEROPS; S07.001; -.
 DR InterPro; IPR001410; DEAD.
 DR InterPro; IPR002464; DEAH_box.
 DR InterPro; IPR001122; Flavi_capsidC.
 DR InterPro; IPR000336; Flavi_glycoproteE.
 DR InterPro; IPR000069; Flavi M.
 DR InterPro; IPR001157; Flavi NS1.
 DR InterPro; IPR000752; Flavi NS2A.
 DR InterPro; IPR000487; Flavi NS2B.
 DR InterPro; IPR000404; Flavi NS4A.
 DR InterPro; IPR001528; Flavi NS4B.
 DR InterPro; IPR000208; Flavi NS5.
 DR InterPro; IPR002535; Flavi_propep.
 DR InterPro; IPR001650; Helicase C.
 DR InterPro; IPR001850; Peptidase S7.
 DR InterPro; IPR007095; RNA_pol_DS_PS.
 DR InterPro; IPR007094; RNA_pol_PSVir.
 DR InterPro; IPR002877; RrmJftsJ_mtfase.
 DR Pfam; PF01003; Flavi capsid; 1.
 DR Pfam; PF02832; Flavi_glycop_C; 1.
 DR Pfam; PF00869; Flavi_glycoprot; 1.
 DR Pfam; PF01004; Flavi M; 1.
 DR Pfam; PF00948; Flavi NS1; 1.
 DR Pfam; PF01005; Flavi NS2A; 1.
 DR Pfam; PF01002; Flavi NS2B; 1.
 DR Pfam; PF01350; Flavi NS4A; 1.
 DR Pfam; PF01349; Flavi NS4B; 1.
 DR Pfam; PF00972; Flavi NS5; 1.
 DR Pfam; PF01570; Flavi_propep; 1.
 DR Pfam; PF01728; FtsJ; 1.
 DR Pfam; PF00271; Helicase C; 1.
 DR Pfam; PF00949; Peptidase S7; 1.
 DR ProDom; PD001556; Flavi_glycoprote; 1.
 DR ProSITE; PS00690; DEAH ATP HELICASE; 1.
 KW ATP-binding; Coat protein; Core protein; Envelope protein;
 KW Glycoprotein; Helicase; Hydrolase; Nonstructural protein; Polyprotein;
 KW RNA-directed RNA polymerase; Transmembrane.
 FT INIT_MET 1 1 Removed from capsid protein C by the
 FT CHAIN 1 112 cellular aminopeptidase.
 FT CHAIN 113 205 Capsid protein C (Potential).
 FT CHAIN 206 280 Envelope protein M (Potential).
 FT CHAIN 281 776 Major envelope protein E (Potential).
 FT CHAIN 777 1128 Nonstructural protein NS1 (Potential).
 FT CHAIN 1129 1358 Nonstructural protein NS2A (Potential).
 FT CHAIN 1359 1489 Flavivirin protease subunit NS2B
 FT CHAIN 1490 2110 (Potential).
 FT CHAIN 2111 2259 Flavivirin protease subunit NS3
 FT CHAIN 2260 2511 (Potential).
 FT CHAIN 2512 3414 Nonstructural protein NS4A (Potential).
 FT NP_BIND 1688 1695 Nonstructural protein NS4B (Potential).
 FT ACT_SITE 1543 1543 RNA-directed RNA polymerase (Potential).
 FT ACT_SITE 1543 1543 ATP (Potential).
 FT ACT_SITE 1567 1567 Charge relay system (By similarity).
 FT ACT_SITE 1627 1627 Charge relay system (By similarity).
 FT SITE 1779 1782 DEAH Box.
 FT TRANSMEM 103 119 Potential.
 FT TRANSMEM 262 278 Potential.
 FT TRANSMEM 728 744 Potential.
 FT TRANSMEM 758 774 Potential.

FT DISULFID 283 310 By similarity.
 FT DISULFID 340 396 By similarity.
 FT DISULFID 354 385 By similarity.
 FT DISULFID 372 401 By similarity.
 FT DISULFID 466 570 By similarity.
 FT CARBOHYD 144 144 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 434 434 N-linked (GlcNAc...) (Potential).
 SQ SEQUENCE 3414 AA; 59CB7E95DD70D82E CRC64;
 Query Match 96.6%; Score 28; DB 1; Length 3414;
 Best Local Similarity 85.7%; Pred. No. 4.2e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 VMXVAEF 7
 DB 929 VMTVAEF 935
 RESULT 3
 Q9IG39 PRELIMINARY; PRT; 3414 AA.
 ID AC Q9IG39;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Polyprotein precursor.
 OS Langkat virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Flavivirus.
 OX NCBI_TaxID=11085;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=attenuated strain E5;
 RX MEDLINE=20192178; PubMed=10725214; DOI=10.1006/viro.2000.0220;
 RA Campbell M.S., Pletnev A.G.;
 RT "Infectious cDNA clones of Langkat tick-borne flavivirus that differ
 RT from their parent in peripheral neurovirulence";
 RL Virology 289:225-237(2000).
 DR EMBL; AF253420; AAF75260.1; -.
 DR HSSP; P14336; 1SVB.
 DR GO; GO:0019028; C:viral capsid; IEA.
 DR GO; GO:0019031; C:viral envelope; IEA.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
 DR GO; GO:0016787; F:hydrolase activity; IEA.
 DR GO; GO:0003723; F:RNA binding; IEA.
 DR GO; GO:0003968; F:RNA helicase activity; IEA.
 DR GO; GO:0003198; F:RNA-directed RNA polymerase activity; IEA.
 DR GO; GO:0019079; F:viral genome replication; IEA.
 DR InterPro; IPR001410; DEAD.
 DR InterPro; IPR011545; DEAD/DEAH_N.
 DR InterPro; IPR001122; Flavi capsidC.
 DR InterPro; IPR011492; Flavi_DEAD.
 DR InterPro; IPR000336; Flavi_glycoproteE.
 DR InterPro; IPR000069; Flavi M.
 DR InterPro; IPR001157; Flavi NS1.
 DR InterPro; IPR000752; Flavi NS2A.
 DR InterPro; IPR000487; Flavi NS2B.
 DR InterPro; IPR000404; Flavi NS4A.
 DR InterPro; IPR001528; Flavi NS4B.
 DR InterPro; IPR000208; Flavi NS5.
 DR InterPro; IPR002535; Flavi_propep.
 DR InterPro; IPR001650; Helicase C.
 DR InterPro; IPR001850; Peptidase S7.
 DR InterPro; IPR007095; RNA_pol_DS_PS.
 DR InterPro; IPR007094; RNA_pol_PSVir.
 DR InterPro; IPR002877; RrmJftsJ_mtfase.
 DR Pfam; PF01003; Flavi capsid; 1.
 DR Pfam; PF02832; Flavi_glycop_C; 1.
 DR Pfam; PF00869; Flavi_glycoprot; 1.
 DR Pfam; PF01004; Flavi M; 1.
 DR Pfam; PF00948; Flavi NS1; 1.
 DR Pfam; PF01005; Flavi NS2A; 1.
 DR Pfam; PF01002; Flavi NS2B; 1.
 DR Pfam; PF01350; Flavi NS4A; 1.
 DR Pfam; PF01349; Flavi NS4B; 1.
 DR Pfam; PF00972; Flavi NS5; 1.
 DR Pfam; PF01570; Flavi_propep; 1.
 DR Pfam; PF01728; FtsJ; 1.
 DR Pfam; PF00271; Helicase C; 1.
 DR Pfam; PF00949; Peptidase S7; 1.
 DR ProDom; PD001556; Flavi_glycoprote; 1.
 DR ProSITE; PS00690; DEAH ATP HELICASE; 1.
 KW ATP-binding; Coat protein; Core protein; Envelope protein;
 KW Glycoprotein; Helicase; Hydrolase; Nonstructural protein; Polyprotein;
 KW RNA-directed RNA polymerase; Transmembrane.
 FT INIT_MET 1 1 Removed from capsid protein C by the
 FT CHAIN 1 112 cellular aminopeptidase.
 FT CHAIN 113 205 Capsid protein C (Potential).
 FT CHAIN 206 280 Envelope protein M (Potential).
 FT CHAIN 281 776 Major envelope protein E (Potential).
 FT CHAIN 777 1128 Nonstructural protein NS1 (Potential).
 FT CHAIN 1129 1358 Nonstructural protein NS2A (Potential).
 FT CHAIN 1359 1489 Flavivirin protease subunit NS2B
 FT CHAIN 1490 2110 (Potential).
 FT CHAIN 2111 2259 Flavivirin protease subunit NS3
 FT CHAIN 2260 2511 (Potential).
 FT CHAIN 2512 3414 Nonstructural protein NS4A (Potential).
 FT NP_BIND 1688 1695 Nonstructural protein NS4B (Potential).
 FT ACT_SITE 1543 1543 RNA-directed RNA polymerase (Potential).
 FT ACT_SITE 1543 1543 ATP (Potential).
 FT ACT_SITE 1567 1567 Charge relay system (By similarity).
 FT ACT_SITE 1627 1627 Charge relay system (By similarity).
 FT SITE 1779 1782 DEAH Box.
 FT TRANSMEM 103 119 Potential.
 FT TRANSMEM 262 278 Potential.
 FT TRANSMEM 728 744 Potential.
 FT TRANSMEM 758 774 Potential.

```

DR Pfam; PF01004; Flavi_M; 1.
DR Pfam; PF00948; Flavi_NS1; 1.
DR Pfam; PF01005; Flavi_NS2A; 1.
DR Pfam; PF01002; Flavi_NS2B; 1.
DR Pfam; PF01350; Flavi_NS4A; 1.
DR Pfam; PF01349; Flavi_NS4B; 1.
DR Pfam; PF00972; Flavi_NS5; 1.
DR Pfam; PF01570; Flavi_Prop; 1.
DR Pfam; PF01728; FtsJ; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR Pfam; PF00949; Peptidase_S7; 1.
DR Pfam; PF01556; Flavi_GlycoproteE; 1.
DR ProDom; PD001496; Flavi_NS1; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELIC; 1.
KW ATP-binding; Helicase; Hydrolase; Polyprotein; Signal.
FT SIGNAL 97 116 Potential.
FT CHAIN 780 1128 nonstructural protein NS1.
FT CHAIN 1129 1358 nonstructural protein NS2a.
FT CHAIN 1359 1489 nonstructural protein NS2b.
FT CHAIN 1490 2110 nonstructural protein NS3.
FT CHAIN 2111 2259 nonstructural protein NS4a.
FT CHAIN 2260 2511 nonstructural protein NS4b.
FT CHAIN 117 280 membrane-associated glycoprotein precursor.
FT CHAIN 2512 3414 nonstructural protein NS5.
FT CHAIN 281 779 envelope membrane-associated glycoprotein.
FT CHAIN 1 96 capsid protein.
SQ SEQUENCE 3414 AA; 378018 MW; BF8F8ACEAB96D534 CRC64;

Query Match 96.6%; Score 28; DB 2; Length 3414;
Best Local Similarity 85.7%; Pred. No. 4.2e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
DB 929 VMTVAEF 935

RESULT 4
Q9IG40 PRELIMINARY; PRT; 3414 AA.
AC Q9IG40;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polyprotein precursor.
OS Langkat virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Flavivirus.
OX NCBI_TaxID=11085;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TP21;
RX MEDLINE=201925178; PubMed=10725214; DOI=10.1006/viro.2000.0220;
RA Campbell M.S., Pletnev A.G.;
RT "Infectious cDNA clones of Langkat tick-borne flavivirus that differ from their parent in peripheral neurovirulence.";
RL Virology 269:225-237(2000).
DR EMBL; AF253419; AAF75259.1; -.
DR PIR; B41704; B41704.
DR HSP; P14336; 1SVB.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003724; F:RNA helicase activity; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.

```

```

DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR001122; Flavi_CapsidC.
DR InterPro; IPR011492; Flavi_DEAD.
DR InterPro; IPR000336; Flavi_GlycoproteE.
DR InterPro; IPR000069; Flavi_M.
DR InterPro; IPR001157; Flavi_NS1.
DR InterPro; IPR000752; Flavi_NS2A.
DR InterPro; IPR000487; Flavi_NS2B.
DR InterPro; IPR000404; Flavi_NS4A.
DR InterPro; IPR001528; Flavi_NS4B.
DR InterPro; IPR000208; Flavi_NS5.
DR InterPro; IPR002535; Flavi_Prop.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR001850; Peptidase_S7.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR InterPro; IPR002877; RnmJftsJ_mtrase.
DR Pfam; PF01003; Flavi_Capsid; 1.
DR Pfam; PF07652; Flavi_DEAD; 1.
DR Pfam; PF00869; Flavi_Glycoprote; 1.
DR Pfam; PF02832; Flavi_Glycop_C; 1.
DR Pfam; PF01004; Flavi_M; 1.
DR Pfam; PF00948; Flavi_NS1; 1.
DR Pfam; PF01005; Flavi_NS2A; 1.
DR Pfam; PF01002; Flavi_NS2B; 1.
DR Pfam; PF01350; Flavi_NS4A; 1.
DR Pfam; PF01349; Flavi_NS4B; 1.
DR Pfam; PF00972; Flavi_NS5; 1.
DR Pfam; PF01570; Flavi_Prop; 1.
DR Pfam; PF01728; FtsJ; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR Pfam; PF00949; Peptidase_S7; 1.
DR ProDom; PD001556; Flavi_GlycoproteE; 1.
DR ProDom; PD001496; Flavi_NS1; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELIC; 1.
KW ATP-binding; Helicase; Hydrolase; Polyprotein; Signal.
FT SIGNAL 97 116 Potential.
FT CHAIN 780 1128 nonstructural protein NS1.
FT CHAIN 1129 1358 nonstructural protein NS2a.
FT CHAIN 1359 1489 nonstructural protein NS2b.
FT CHAIN 1490 2110 nonstructural protein NS3.
FT CHAIN 2111 2259 nonstructural protein NS4a.
FT CHAIN 2260 2511 nonstructural protein NS4b.
FT CHAIN 117 280 membrane-associated glycoprotein precursor.
FT CHAIN 2512 3414 nonstructural protein NS5.
FT CHAIN 281 779 envelope membrane-associated glycoprotein.
FT CHAIN 1 96 capsid protein.
SQ SEQUENCE 3414 AA; 378138 MW; 8DE86A4A0E8F7E3 CRC64;

Query Match 96.6%; Score 28; DB 2; Length 3414;
Best Local Similarity 85.7%; Pred. No. 4.2e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
DB 929 VMTVAEF 935

RESULT 5
Q7XA44 PRELIMINARY; PRT; 433 AA.
AC Q7XA44;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Putative transposase-related protein.
GN ORFNames=CB3A14.3;
OS Solanum bulbocastanum (Wild potato).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta;

```

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asteride;
 OC Lamiales; Solanales; Solanaceae; Solanum.
 OX NCBI_TaxID=147425;
 RN
 RN SEQUENCE FROM N.A.
 RA PubMed=12872003; DOI=10.1073/pnas.1533501100;
 RA Song J., Bradeen J.M., Naess S.K., Raasch J.A., Wielgus S.M.,
 RA Haberglach G.T., Liu J., Kuang H., Austin-Phillips S., Buell C.R.,
 RA Helgeson J.P., Jiang J.;
 RA "Gene RB cloned from Solanum bulbocastanum confers broad spectrum
 RT resistance to potato late blight."
 RL Proc. Natl. Acad. Sci. U.S.A. 100:9128-9133 (2003).
 DR EMBL: AY303170; AAP45161.1; -;
 DR InterPro: IPR000817; Prion.
 SQ SEQUENCE 433 AA; 49486 MW; 9AE824BB05D384C1 CRC64;
 Query Match 93.1%; Score 27; DB 2; Length 433;
 Best Local Similarity 71.4%; Pred. No. 1.le+02;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 VMKVAEP 7
 Db : |||||
 29 IMTVAEP 35
 RESULT 6
 POT3 ARATH STANDARD; PRT; 775 AA.
 ID POT3 ARATH STANDARD; PRT; 775 AA.
 AC Q9FE38; O22399; Q9SUR2;
 DT 29-MAR-2004 (Rel. 43, Created)
 DT 29-MAR-2004 (Rel. 43, Last sequence update)
 DT 03-JUL-2004 (Rel. 44, Last annotation update)
 DE Potassium transporter 3 (AtPOT3) (AtKUP4) (AtKT3) (Tiny root hair 1
 DE protein).
 GN Name:POT3; Synonyms:KT3, KUP4, TRH1; OrderedLocusNames=At4g23640;
 GN ORFNames=F9D16.110;
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosid II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RN SEQUENCE FROM N.A., FUNCTION, AND NULL MUTANT.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=21096391; PubMed=11158535;
 RA Rigas S., Debrosses G., Haralampidis K., Vicente-Agullo F.,
 RA Feldmann K.A., Grabov A., Dolan L., Hatzopoulos P.;
 RT "TRH1 encodes a potassium transporter required for tip growth in
 RT Arabidopsis root hairs."
 RL Plant Cell 13:139-151 (2001).
 RN [2]
 RN SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=20083488; PubMed=10617198; DOI=10.1038/47134;
 RA Mayer K.F.X., Schueller C., Wambutt R., Murphy G., Volckaert G.,
 RA Pohl T., Duesterhoeft A., Stiekema W., Entian K.-D., Terryn N.,
 RA Harris B., Ansoorge W., Brandt P., Grivell L.A., Rieger M.,
 RA Weichselgartner M., de Simone V., Obermaier B., Maché R., Mueller M.,
 RA Kreis M., Delseny M., Puigdomenech P., Watson M., Schmidtheini T.,
 RA Reichert B., Portetelie D., Perez-Alonso M., Boutry M., Bancroft I.,
 RA Vos P., Hohnsels J., Zimmermann W., Wedler H., Ridley P.,
 RA Langham S.-A., McCullagh B., Bilham L., Robben J.,
 RA Van der Schueren J., Grynoprez B., Chuang Y.-J., Vandenbussche F.,
 RA Braeken M., Weltjens I., Voet M., Bastiaens I., Aert R., Defoor E.,
 RA Weitzenecker T., Bothe G., Ramsperger U., Hilbert H., Braun M.,
 RA Holzer E., Brandt A., Peters S., van Staveren M., Dirkse W.,
 RA Moolman P., Klein Lankhorst R., Roe M., Hauf J., Koeter P.,
 RA Berner S., Hempel S., Feldpausch M., Lamberth S., Van den Daele H.,
 RA De Keyser A., Buyschaert C., Giehl J., Cronin A., Quail M.A., Bray-Allen S.,
 RA Van Montagu M., Rogers J., Hall S., Kay M., Lennard N., McLay K., Mayes R.,
 RA Clark L., Doggett J., Hall S., Kay M., Lennard N., McLay K., Mayes R.,
 RA Pettett A., Rajandream M.A., Lyne M., Benes V., Rechmann S.,
 RA Borkova D., Bloecker H., Scharfe M., Grimm M., Loehnert T.-H.,

RA Dose S., de Haan M., Maarse A.C., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fartmann B., Granderath K., Dauner D., Herzl A.,
 RA Neumann S., Argiriou A., Vitale D., Liguori R., Piravandi E.,
 RA Massenot O., Quigley F., Clabault G., Muendlein A., Felber R.,
 RA Schnabl S., Hiller R., Schmidt W., Lecharny A., Aubourg S.,
 RA Chedor F., Cooke R., Berger C., Monfort A., Casacuberta E.,
 RA Gibbons T., Weber N., Vandenbol M., Bagues M., Terol J., Torres A.,
 RA Perez-Perez A., Purnelle B., Bent E., Johnson S., Tacon D., Jesse C.,
 RA Heijnen L., Schwarz S., Scholler P., Heber S., Francis P., Biele C.,
 RA Frishman D., Haase D., Lemcke K., Mewes H.-W., Stocker S.,
 RA Zaccaria P., Bevan M., Wilson R.K., de la Bastide M., Habermann K.,
 RA Parnell L., Dedhia N., Gnoj L., Schutz K., Huang E., Spiegel L.,
 RA Sekton M., Murray J., Sheet P., Cordes M., Abu-Threideh J.,
 RA Stoneking T., Kalicki J., Graves T., Harmon G., Edwards J.,
 RA Latreille P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,
 RA Minx P., Bentley D., Fulton B., Miller N., Greco T., Kemp K.,
 RA Kramer J., Fulton L., Mardis E., Dante M., Pepin K., Hillier L.W.,
 RA Nelson J., Spieth J., Ryan E., Andrews S., Geisel C., Layman D.,
 RA Du H., Ali J., Berghoff A., Jones K., Drone K., Cotton M., Joshi C.,
 RA Antonoli B., Zidanic M., Strong C., Sun H., Lamar B., Yordan C.,
 RA Ma P., Zhong J., Preston R., Vill D., Shekher M., Matero A., Shah R.,
 RA Swaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Till S.,
 RA Granat S., Shohdy N., Hasegawa A., Hameed A., Lodhi M., Johnson A.,
 RA Chen E., Marra M.A., Martienssen R., McCombie W.R.;
 RT "Sequence and analysis of chromosome 4 of the plant Arabidopsis
 RT thaliana.";
 RL Nature 402:769-777 (1999).
 RN [3]
 RN SEQUENCE OF 1-344 FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=98010480; PubMed=9350997; DOI=10.1016/S0014-5793 (97)01125-3;
 RA Quintero F.J., Blatt M.R.;
 RT "A new family of K⁺ transporters from Arabidopsis that are conserved
 RT across phyla.";
 RL FEBS Lett. 415:206-211 (1997).
 RN [4]
 RN TISSUE SPECIFICITY
 RX MEDLINE=98138058; PubMed=9477571;
 RA Kim E.J., Kwak J.M., Uozumi N., Schroeder J.I.;
 RT "AtKUP1: an Arabidopsis gene encoding high-affinity potassium
 RT transport activity.";
 RL Plant Cell 10:51-62 (1998).
 RN [5]
 RN GENE FAMILY, AND NOMENCLATURE.
 RX MEDLINE=21392307; PubMed=11500563; DOI=10.1104/pp.126.4.1846;
 RA Maeser P., Thominé S., Schroeder J.I., Ward J.M., Hirschi K., Sze H.,
 RA Talke I.N., Amtmann A., Maathuis F.J.M., Sanders D., Harper J.F.,
 RA Tchieu J., Gribskov M., Persans M.W., Salt D.E., Kim S.A.,
 RA Gueriot M.L.;
 RT "Phylogenetic relationships within cation transporter families of
 RT Arabidopsis.";
 RL Plant Physiol. 126:1646-1667 (2001).
 CC -!- FUNCTION: High-affinity potassium transporter required for tip
 CC growth of root hairs.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Plasma membrane
 CC (Potential).
 CC -!- TISSUE SPECIFICITY: Detected at very low levels in roots, stems,
 CC leaves and flowers of mature plants.
 CC -!- DEVELOPMENTAL STAGE: Highly expressed in roots in the early stage
 CC of seedlings growth.
 CC -!- MISCELLANEOUS: Loss-of-function mutation results in the arrest of
 CC root hair growth.
 CC -!- SIMILARITY: Belongs to the HAK/KUP transporter (TC 2.A.72) family.
 CC -!- CAUTION: Ref.2 sequences differ from that shown due to erroneous
 CC gene model prediction.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>).


```

CC -----
DR ENBL; AJ296155; CAC16137.1; -.
DR EMBL; AJ296156; CAC16138.1; -.
DR EMBL; AL035394; CAA23030.1; ALT_SEQ.
DR EMBL; AL161559; CAB79319.1; ALT_SEQ.
DR ENBL; AF012858; AAC49846.1; -.
DR PIR; T05596; T05596.
DR InterPro; IPR003855; K+ transporter.
DR TIGRFAMs; TIGR00794; Kup; 1.
DR Pfam; PF02705; K trans; 1.
KW Multigene family; Potassium transport; Transmembrane; Transport.
FT DOMAIN 1 10 Cytoplasmic (Potential).
FT TRANSMEM 11 31 Potential.
FT DOMAIN 32 54 Extracellular (Potential).
FT TRANSMEM 55 75 Potential.
FT DOMAIN 76 143 Cytoplasmic (Potential).
FT TRANSMEM 144 164 Potential.
FT DOMAIN 165 177 Extracellular (Potential).
FT TRANSMEM 178 198 Potential.
FT DOMAIN 199 205 Cytoplasmic (Potential).
FT TRANSMEM 206 226 Potential.
FT DOMAIN 227 228 Extracellular (Potential).
FT TRANSMEM 229 249 Potential.
FT DOMAIN 250 253 Cytoplasmic (Potential).
FT TRANSMEM 254 274 Potential.
FT DOMAIN 275 286 Extracellular (Potential).
FT TRANSMEM 287 307 Potential.
FT DOMAIN 308 327 Cytoplasmic (Potential).
FT TRANSMEM 328 348 Potential.
FT DOMAIN 349 378 Extracellular (Potential).
FT TRANSMEM 379 399 Potential.
FT DOMAIN 400 407 Cytoplasmic (Potential).
FT TRANSMEM 408 428 Potential.
FT DOMAIN 429 435 Extracellular (Potential).
FT TRANSMEM 436 456 Potential.
FT DOMAIN 457 463 Cytoplasmic (Potential).
FT TRANSMEM 464 484 Potential.
FT DOMAIN 485 775 Extracellular (Potential).
SQ SEQUENCE 775 AA; B0C55068B4BE180 CRC64;

Query Match 93.1%; Score 27; DB 1; Length 775;
Best Local Similarity 71.4%; Pred. No. 2e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
DB 602 VMSIAEF 608
-----
RESULT 7
ID QBVXQ3 PRELIMINARY; PRT; 814 AA.
AC QBVXQ3;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Putative potassium transporter.
GN Names-hak2;
OS Cymodocea nodosa.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Cymodoceaceae; Cymodocea.
OX NCBI_TaxID=55448;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22261168; PubMed=12374286; DOI=10.1023/A:1019951023362;
RA Garcadeblas B., Benito B., Rodriguez-Navarro A.;
RT "Molecular cloning and functional expression in bacteria of the
RT potassium transporters CnHAK1 and CnHAK2 of the seagrass Cymodocea
RT nodosa.";
RL Plant Mol. Biol. 50:623-633 (2002).
DR ENBL; AJ427293; CAD20319.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0015079; F:potassium ion transporter activity; IEA.

```

```

DR GO; GO:0006813; P:potassium ion transport; IEA.
DR Pfam; PF02705; K trans; 1.
DR TIGRFAMs; TIGR00794; kup; 1.
SQ SEQUENCE 814 AA; 91608 MW; 6350453B857BECB CRC64;

Query Match 93.1%; Score 27; DB 2; Length 814;
Best Local Similarity 71.4%; Pred. No. 2.1e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
DB 611 VMSIAEF 617
-----
RESULT 8
ID DEOC_MYCGE STANDARD; PRT; 223 AA.
AC P47296;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Deoxyribose-phosphate aldolase (EC 4.1.2.4) (Phosphodeoxyriboaldolase)
DE (Deoxyriboaldolase) (DERA).
GN Names-deoc; OrderedLocustNames=MG050;
OS Mycoplasma genitalium.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2097;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 33530 / G-37;
RX MEDLINE=96026346; PubMed=7569993;
RA Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,
RA Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G.G., Kelley J.M.,
RA Fritchman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhmann J.L.,
RA Nguyen D.T., Utterback T.R., Saudek D.M., Phillips C.A., Merrick J.M.,
RA Tomb J.-F., Dougherty B.A., Bort K.F., Hu P.-C., Lucier T.S.,
RA Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;
RT "The minimal gene complement of Mycoplasma genitalium.";
RL Science 270:397-403 (1995).
CC -|- CATALYTIC ACTIVITY: 2-deoxy-D-ribose 5-phosphate = D-
CC glyceraldehyde 3-phosphate + acetaldehyde.
CC -|- PATHWAY: Nucleotide and deoxyribonucleotide catabolism.
CC -|- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -|- SIMILARITY: Belongs to the deoc/fbaB aldolase family. Deoc
CC subfamily 1.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U39684; AAC71266.1; -.
CC PIR; E64205; E64205.
CC HSSP; O66540; 1MZH.
CC TIGR; MG050; -.
CC HAMAP; MF_00114; -.
CC InterPro; IPR002915; Deoc.
CC InterPro; IPR003009; FMN_enzyme.
CC Pfam; PF01791; Deoc; 1.
KW Complete proteome; Lyase; Schiff base.
FT BINDING 152 152 Schiff-base with acetaldehyde (By
FT similarity).
FT ACT SITE 181 181 By similarity.
SQ SEQUENCE 223 AA; 24675 MW; 332430231CE99DB0 CRC64;

Query Match 89.7%; Score 26; DB 1; Length 223;
Best Local Similarity 71.4%; Pred. No. 1.1e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7

```

Db 93 VMNIAEF 99
||:||||

RESULT 9

DEOC_MYCPN	STANDARD;	PRT;	224 AA.
ID _DEOC_MYCPN			
AC	P09924;		
DT	01-MAR-1989 (Rel. 10, Created)		
DT	01-MAR-1989 (Rel. 10, Last sequence update)		
DT	05-JUL-2004 (Rel. 44, Last annotation update)		
DE	Deoxyribose-phosphate aldolase (EC 4.1.2.4) (Phosphodeoxyribosaldolase)		
DE	(Deoxyribosaldolase) (DERA).		
GN	Name=deoc; OrderedLocusNames=MPN063; ORFNames=MP091;		
OS	Mycoplasma pneumoniae.		
OC	Bacteria; Firmicutes; Mollicutes; Mycoplastmataceae; Mycoplasma.		
NCBI_TaxID=2104;			
[1]			
RP	SEQUENCE FROM N.A.		
RC	STRAIN=ATCC 29342 / M129;		
RA	MEDLINE=89128453; PubMed=2492658;		
RX	Loechele S., Inamine J.M., Hu P.-C.;		
RT	Nucleotide sequence of the deoC gene of Mycoplasma pneumoniae.;		
RL	Nucleic Acids Res 17:801-801(1989).		

RP SEQUENCE FROM N.A.

RC STRAIN=ATCC 29342 / M29;
RX MEDLINE=97105885; PubMed=8948633; DOI=10.1093/nar/24.22.4420;
RA Hamelreich R., Hilbert H., Plagens H., Pirkel E., Li B.-C.,
RA Herrmann R.;
RT "Complete sequence analysis of the genome of the bacterium *Mycoplasma*
RT *pneumoniae*."; Nucleic Acids Res. 24:4420-4449(1996).
RL -I- CATALYTIC ACTIVITY: 2'-deoxy-D-ribose 5-phosphate = D-
CC glyceraldehyde 3-phosphate + acetaldehyde.
CC -I- PATHWAY: Nucleotide and deoxyribonucleotide catabolism.
CC -I- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -I- SIMILARITY: Belongs to the deoc/fbaB aldolase family. Deoc
C subfamily 1.

 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).

DR EMBL: X13544: CAA3

DR	EMBL; AB03941; CA838277.1; -.
DR	EMBL; AE000011; AAB95739.1; -.
DR	PIR; S02216; S02216.
DR	HSSP; Q9X1P5; 100Y.
DR	HMAP; MF_00114; -, 1.
DR	InterPro; IPR002915; DeoC.
DR	Pfam; PF01791; DeoC; 1.
KW	Complete proteome; Lyase; Schiff base.
FT	BINDING 152 152 Schiff-base with acetaldehyde (By similarity).
FT	ACT SITE 181 181 By similarity.
SO	SEQUENCE 224 AA; 24878 MW; 73C3E4932E7881F7 CRC64;

Query Match	89.7%	Score 26;	DB 1;	Length 224;
Best Local Similarity	71.4%	Pred. No. 1.2e+02;		
Matches	5.	Conservative	1.	Mismatches 0.
				Indels 0.

02 1 VMXVAF 7

QY 1 VMXVAF 7
||: |||
Db 93 VMNIAEF 99

RESULT 10

RESULTS TO
Q8U0A9
ID Q8U0A9 PRELIMINARY; PRT; 262 AA.

08110A9:

AC	Q0093;	
AD	01-JUN-2002 (T:EMBLrel. 21, Created)	
AE	01-JUN-2002 (T:EMBLrel. 21, Last sequence update)	
AF	01-MAR-2002 (T:EMBLrel. 26, Last annotation update)	
AG	2-dehydro-3-deoxyphosphoheptonate aldolase (EC 4.1.2.15).	
AN	OrderedLocusNames=PF1690;	
AO	Pyrococcus furiosus.	
AP	Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;	
AR	Pyrococcus.	
AS	NCBI_TaxID=2261;	
AT	[1]	
AV	SEQUENCE FROM N.A.	
AW	STRAIN=Vc1 / DSM 3638 / ATCC 43587 / JCM 8422;	
AX	Weiss R.B., Dunn D.M., Robb P.T., Brown J.R.;	
AY	"The complete sequence of the Pyrococcus furiosus genome."	
AZ	Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.	
BA	EMBL; AE010267; AAL81814.1; -.	
BB	HSP; O66496; 1PE1.	
BC	GO; GO:0016832; F:aldehyde-lyase activity; IEA.	
BD	GO; GO:0009073; P:aromatic amino acid family biosynthesis; IEA.	
BE	GO; GO:0009058; P:biosynthesis; IEA.	
BF	Pfam; PF00793; DAHP_synth_1; 1.	
BG	TIGRfams; TIGR01361; DAHP_synth_Bsub; 1.	
BH	Complete proteome.	
BI	SEQUENCE 262 AA: 29225 MW: ED84D28FD5DEADC2 CRC64;	

Query Match 89.7%; Score 26; DB 2; Length 262;
Best Local Similarity 71.4%; Pred. No. 1.3e+02;

```

Qy      1 VMXVAEF 7
          : | | | |
Db      39 IMKVAEF 45
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

RESULT 11

Q9ZPQ5 PRELIMINARY; PRT: 324 AA.

AC Q9ZPQ5; 01-MAY-1999 (TrEMBLrel. 10, Created)

DT 01-MAY-1999 (TrEMBLrel. 10, last sequence update)

DT 01-MAR-2004 (TrEMBLrel. 26, last annotation update)

DE Putative steroid sulfotransferase.

DE Name=At2g03770;

GN Arabidopsis thaliana (Mouse-ear cress).

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;

OC eurosid II; Brassicales; Brassicaceae; Arabidopsids.

OC NCBI_TaxID=3702;

EN [1]

FP SEQUENCE FROM N.A.

RP Lin X., Kaul S., Shea T.P., Fujii C.Y., Shen M., VanAken S.E.,

RA Barnstead M.E., Mason T.M., Bowman C.L., Ronning C.M., Benito M.-I.,

RA Carrara A.J., Creasy T.H., Buell C.R., Town C.D., Nierman W.C.,

RA Fraser C.M., Venter J.C.;

RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

RP SEQUENCE FROM N.A.

RA Town C.D., Kaul S.;
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 RL EMBL; AC006836; AAD20079.1; -.
 DR PIR; B84452; B84452.
 DR HSRP; P52839; IQ44.
 DR GO; GO:0008146; F:sulfotransferase activity; IEA.
 DR GO; GO:016740; F:transferase activity; IEA.
 DR InterPro; IPR000863; Sulfotransferase.
 DR Pfam; PF00685; Sulfotransferase; 1.
 DR ProDom; PD001218; Sulfotransferase; 1.
 KW Transferase.

Query Match 89.7%; Score 26; DB 2; Length 324;
Best Local Similarity 71.4%; Pred. No. 1.6e+02;

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEP 7

Db 233 VMRIAEP 239

RESULT 12

ID Q65E92 PRELIMINARY; PRT; 479 AA.

AC Q65E92; 2004 (TReMBLrel. 28, Created)

DT 25-OCT-2004 (TReMBLrel. 28, Last sequence update)

DT 25-OCT-2004 (TReMBLrel. 28, Last annotation update)

DE Hypothetical protein tuab.

GN Name: tuab; ORFNames=BL03350, BL103806;

OS Bacillus licheniformis DSM 13

OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.

OX NCBI_TaxID=279010;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=DSM 13;

RX PubMed=15383718;

RA Veith B., Herzberg C., Steckel S., Feesche J., Maurer K.H.,

RA Ehrenreich P., Baeumer S., Henne A., Liesegang H., Merkl R.,

RA Ehrenreich A., Gottschalk G.;

RT "The Complete Genome Sequence of Bacillus licheniformis DSM13, an

RT Organism with Great Industrial Potential";

RL J. Mol. Microbiol. Biotechnol. 7:204-211(2004).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=ATCC 14580;

RA Rev M.W., Ramaiya P., Nelson B.A., Brody-Karlin S.D., Zaretsky E.J.,

RA Tang M., de Leon A.L., Xiang H., Gusti V., Clausen I.G., Olsen P.B.,

RA Rasmussen M.D., Andersen J.I., Jorgensen P.L., Larsen T.S.,

RA Sorokin A., Bolotin A., Lapidus A., Galleron N., Ehrlich S.D.,

RA Berka R.M.;

RT "Complete genome sequence of the industrial bacterium Bacillus

RT licheniformis and comparisons with closely related Bacillus species.";

RL Genome Biol. 5:R77-R77(2004).

DR ENBL; AE017333; AAU42622.1; -.

DR ENBL; CP000002; AAU25249.1; -.

KW Hypothetical protein.

SQ SEQUENCE 479 AA; 53318 MW; B25425B96AD39D62 CRC64;

Query Match

Best Local Similarity 89.7%; Score 26; DB 2; Length 479;

Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEP 7

Db 36 IMTIAEP 42

RESULT 13

Q6Z5S4

ID Q6Z5S4 PRELIMINARY; PRT; 731 AA.

AC Q6Z5S4;

DT 05-JUL-2004 (TReMBLrel. 27, Created)

DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)

DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)

DE Hypothetical protein OSJNBa0033J14.4.

GN Name=OSJNBa0033J14.4;

OS Oryza sativa (japonica cultivar-group).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

OC Ehrhartoideae; Oryzaceae; Oryza.

OX NCBI_TaxID=39947;

RN [1]

RP SEQUENCE FROM N.A.

RA Sasaki T., Matsumoto T., Katayose Y.;

RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 7, BAC

clone:OSJNBa0033J14.";

RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.

DR ENBL; AP005102; BAC83706.1; -.

DR InterPro; IPR007658; DUF594.

DR Pfam; PF04578; DUF594; 1.

KW Hypothetical protein.

SQ SEQUENCE 731 AA; 82975 MW; 62C331F04D1716AF CRC64;

Query Match

Best Local Similarity 89.7%; Score 26; DB 2; Length 731;

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEP 7

Db 391 IMGVAEP 397

RESULT 14

Q9XBW4

ID Q9XBW4 PRELIMINARY; PRT; 821 AA.

AC Q9XBW4; Q7BWB5;

DT 01-NOV-1999 (TReMBLrel. 12, Created)

DT 01-NOV-1999 (TReMBLrel. 12, Last sequence update)

DT 25-OCT-2004 (TReMBLrel. 28, Last annotation update)

DE Immunoreactive 92 kDa antigen PG21 (Zinc carboxypeptidase,

DE putative).

GN OrderedLocusNames=PG0232;

OS Porphyromonas gingivalis (Bacteroides gingivalis).

OC Bacteria; Bacteroidetes; Bacteroides (Class); Bacteroidales;

OC Porphyromonadaceae; Porphyromonas.

OX NCBI_TaxID=837;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=W50;

RA Ross B.C., Barr I., Patterson M., Agius C., Rothel L., Margetts M.,

RA Hocking D., Webb E.;

RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=W83;

RX MEDLINE=22829867; PubMed=12949112;

RX DOI=10.1128/JB.185.18.5591-5601.2003;

RA Nelson K.E., Fleischmann R.D., DeBoy R.T., Paulsen I.T., Fouts D.E.,

RA Eisen J.A., Dougherty S.C., Dodson R.J., Durkin A.S., Gwinn M.L.,

RA Haft D.H., Kolonay J.F., Nelson W.C., Mason T.M., Tallon L., Gray J.,

RA Granger D., Tettelin H., Dong H., Galvin J.L., Duncan M.J.,

RA Dewhirst F.E., Fraser C.M.;

RT "Complete genome sequence of the oral pathogenic bacterium

RT Porphyromonas gingivalis strain W83.";

RL J. Bacteriol. 185:5591-5601(2003).

DR ENBL; AF153768; AAD38980.1; -.

DR ENBL; AE017172; AAQ65462.1; -.

DR HSSP; Q90240; IQMU.

DR MEROPS; M14.023; -.

DR TIGR; PG0232; -.

DR GO; GO:0004182; F:carboxypeptidase A activity; IEA.

DR GO; GO:0008270; F:zinc ion binding; IEA.

DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.

DR InterPro; IPR008969; CarboxypepD reg.

DR InterPro; IPR000834; Peptidase_M14.

DR InterPro; IPR000601; PKD.

DR Pfam; PF00246; Peptidase_M14; 1.

DR Pfam; PF00801; PKD; 3.

DR PRINTS; PR00765; CRBOXYPTASEA.

DR SMART; SM00089; PKD; 3.

DR SMART; SM00631; Zn pept; 1.

DR PROSITE; PS00133; CARBOXYPEPT_ZN_2; UNKNOWN_1.

DR PROSITE; PS50093; PKD; 3.

KW Complete proteome.

SQ SEQUENCE 821 AA; 91517 MW; 250843B2C9C833E2 CRC64;

Query Match

Best Local Similarity 89.7%; Score 26; DB 2; Length 821;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY      1 VMXVAEF 7
Db      565 VMPVAEF 571

RESULT 15
Q62IK2  PRELIMINARY;      PRT;   205 AA.
AC      Q62IK2;
DT      25-OCT-2004 (T-EMBLrel. 28, Created)
DT      25-OCT-2004 (T-EMBLrel. 28, Last sequence update)
DT      25-OCT-2004 (T-EMBLrel. 28, Last annotation update)
DE      Hypothetical protein.
GN      ORFNames=BMA1869;
OS      Burkholderia mallei ATCC 23344.
OC      Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC      Burkholderiaceae; Burkholderia.
OX      NCBI_TaxID=243160;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=ATCC 23344;
RA      Niernan W.C., DeShazer D., Kim H.S., Tettelin H., Nelson K.E.,
RA      Feldblyum T., Ulrich R.L., Ronning C.M., Brinkac L.M., Daugherty S.C.,
RA      Davidson T.D., DeBoy R.T., Dimitrov G., Dodson R.J., Durkin A.S.,
RA      Gwinn M.L., Haft D.H., Khouri H., Kolonay J.F., Madupu R.,
RA      Mohammed Y., Nelson W.C., Radune D., Romero C.M., Sarrisa S.,
RA      Selengut J., Shamblin C., Sullivan S.A., White O., Yu Y., Zafar N.,
RA      Zhou L., Fraser C.M.;
RT      "Structural flexibility in the Burkholderia mallei genome.";
RL      Proc. Natl. Acad. Sci. U.S.A. 101:14247-14251(2004).
DR      EMBL; CP000010; AAU49424.1; -.
KW      Hypothetical protein.
SQ      SEQUENCE 205 AA; 21152 MW; 67216B40EBF4D59F CRC64;

Query Match      86.2%; Score 25; DB 2; Length 205;
Best Local Similarity 71.4%; Pred. No. 2e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 VMXVAEF 7
Db      55 VLAFAEF 61

RESULT 16
Q63VR6  PRELIMINARY;      PRT;   205 AA.
AC      Q63VR6;
DT      25-OCT-2004 (T-EMBLrel. 28, Created)
DT      25-OCT-2004 (T-EMBLrel. 28, Last sequence update)
DT      25-OCT-2004 (T-EMBLrel. 28, Last annotation update)
DE      Putative membrane protein.
GN      ORFNames=BPSL1178;
OS      Burkholderia pseudomallei K96243.
OC      Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC      Burkholderiaceae; Burkholderia.
OX      NCBI_TaxID=272560;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=K96243;
RA      Holden M.T.G., Titball R.W., Peacock S.J., Cerdano-Tarraga A.M.,
RA      Atkins T., Crossman L.C., Pitt T., Churcher C., Mungall K.,
RA      Bentley S.D., Sebahia M., Thomson N.R., Bason N., Beacham I.R.,
RA      Brooks K., Brown K.A., Brown N.F., Challis G.L., Cherevach I.,
RA      Chillingworth T., Cronin A., Cresset B., Davis P., DeShazer D.,
RA      Felwell T., Fraser A., Hance Z., Hauser H., Holtroyd S., Jagels K.,
RA      Keith K.E., Maddison M., Moule S., Price C., Quail M.A.,
RA      Rabinowitz E., Rutherford K., Sanders M., Simmonds M.,
RA      Songvilai S., Stevens K., Tumapa S., Vesaratchaveit M.,
RA      Whitehead S., Yeats C., Barrall B.G., Oyston P.C.F., Parkhill J.;
RT      "Genomic plasticity of the causative agent of melioidosis,
RT      Burkholderia pseudomallei.";
```

```
RL      Proc. Natl. Acad. Sci. U.S.A. 101:14240-14245 (2004).
DR      EMBL; BX571965; CAH35173.1; -. 67216B40EBF4D59F CRC64;
SQ      SEQUENCE 205 AA; 21152 MW; 67216B40EBF4D59F CRC64;

Query Match      86.2%; Score 25; DB 2; Length 205;
Best Local Similarity 71.4%; Pred. No. 2e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 VMXVAEF 7
Db      55 VLAFAEF 61

RESULT 17
YTBO_BACSU  STANDARD;      PRT;   253 AA.
AC      P53560; O34914;
DT      01-OCT-1996 (Rel. 34, Created)
DT      01-OCT-1996 (Rel. 34, Last sequence update)
DT      25-OCT-2004 (Rel. 45, Last annotation update)
DE      Hypothetical protein ytbQ.
GN      Name=ytbQ; OrderedLocusNames=BSU30180;
OS      Bacillus subtilis.
OC      Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX      NCBI_TaxID=1423;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      MEDLINE=96312354; PubMed=8763940;
RA      Bower S., Perkins J.B., Yocum R.R., Howitt C.L., Rahaim P., Pero J.;
RT      "Cloning, sequencing, and characterization of the Bacillus subtilis
RT      biotin biosynthetic operon.";
RL      J. Bacteriol. 178:4122-4130(1996).
RN      [2]
RP      SEQUENCE FROM N.A.
RA      Lapidus A., Galleron N., Sorokin A., Ehrlich S.D.;
RL      Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
RN      [3]
RP      SEQUENCE FROM N.A.
RC      STRAIN=168;
RX      MEDLINE=98044033; PubMed=9384377; DOI=10.1038/36786;
RA      Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA      Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA      Borriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA      Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA      Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA      Denizot F., Devine K.M., Dusterhoft A., Ferrari E., Foulger D.,
RA      Entian K.-D., Errington J., Fabret C., Ehrlich S.D., Emerson P.T.,
RA      Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA      Chim S.Y., Glaeser P., Goffeau A., Golightly E.J., Grandi G.,
RA      Guiseppe G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,
RA      Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M.,
RA      Jones L.-M., Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M.,
RA      Klein C., Kobayashi Y., Koetter P., Koningsstein G., Krogh S.,
RA      Kumano M., Kurita K., Lapidus A., Lardinois S., Lauber J.,
RA      Lazarevic V., Lee S.M., Levine A., Liu H., Masuda S., Mauel C.,
RA      Medigue C., Medina N., Mellado R.P., Mizuno M., Moesl D., Nakai S.,
RA      Noback M., Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B.,
RA      Park S.H., Parro V., Pohl T.M., Portetelle D., Porwollik S.,
RA      Prescott A.M., Prescan E., Pujic P., Purnelle B., Rapoport G.,
RA      Rey M., Reynolds S., Rieger M., Rivolta C., Rocha E., Roche B.,
RA      Rose M., Sadale Y., Sato T., Scanlan E., Schleich S., Schroeter R.,
RA      Scoffone F., Sekiguchi J., Sekowska A., Seror S.J., Seror P.,
RA      Shin B.S., Soldo B., Sorokin A., Tacconi E., Takagi T., Takahashi H.,
RA      Takenaru K., Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P.,
RA      Tognoni A., Tosato V., Uchiyama S., Vandenbol M., Vannier F.,
RA      Vassarotti A., Viari A., Wambutt R., Wedler E., Wedler H.,
RA      Weitzengesser T., Winters P., Wipat A., Yamamoto H., Yamane K.,
RA      Yasumoto K., Yata K., Yoshida K., Yoshikawa H.F., Zumstein E.,
RA      Yoshikawa H., Danchin A.;
RT      "The complete genome sequence of the Gram-positive bacterium Bacillus
RT      subtilis.";
RL      Nature 390:249-256(1997).
RC      -----
```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)
 CC or send an email to license@isb-sib.ch)

DR EMBL; U51868; BAB17463.1; -;
 DR EMBL; AF008220; AAC00267.1; ALT INIT.
 DR EMBL; Z99119; CAB14996.1; ALT_INIT.
 DR Subtilisin; BGL1787; ybQ.
 KW Complete proteome; Hypothetical protein.
 SQ SEQUENCE 253 AA; 28233 MW; 572CC4681F04C650 CRC64;

Query Match 86.2%; Score 25; DB 1; Length 253;
 Best Local Similarity 57.1%; Pred. No. 2.4e+02;
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
 : : : : :
 Db 75 IMDIAPF 81

RESULT 18
 Q8U3Z7 PRELIMINARY; PRT; 284 AA.
 ID Q8U3Z7; DSM 3638 / ATCC 43587 / JCM 8422;
 AC Q8U3Z7; Weiss R.B., Dunn D.M., Robb F.T., Brown J.R.;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Co-factor modifying protein.
 GN OrderedLocusNames=PF0302;
 OS Pyrococcus furiosus.
 OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
 OC Pyrococcus.
 OX NCBI_TaxID=2261;

SEQUENCE FROM N.A.
 RC STRAIN=Vc1 / DSM 3638 / ATCC 43587 / JCM 8422;
 RA Weiss R.B., Dunn D.M., Robb F.T., Brown J.R.;
 RT "The complete sequence of the Pyrococcus furiosus genome."
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AE010154; AAL80426.1; -;
 DR GO; GO:0003824; F:catalytic activity; IEA.
 DR GO; GO:0005506; F:iron ion binding; IEA.
 DR Pfam; PF04055; Radical SAM; 1.
 DR SMART; SM00729; Elp3; 1.
 KW Complete proteome.
 SQ SEQUENCE 284 AA; 32512 MW; B19557B14451FFCD CRC64;

Query Match 86.2%; Score 25; DB 2; Length 284;
 Best Local Similarity 57.1%; Pred. No. 2.7e+02;
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
 : : : : :
 Db 222 IMKIAEF 228

RESULT 19
 Q82ZH7 PRELIMINARY; PRT; 406 AA.
 ID Q82ZH7; DSM 3638 / ATCC 43587 / JCM 8422;
 AC Q82ZH7; Weiss R.B., Dunn D.M., Robb F.T., Brown J.R.;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Peptidase T.
 GN Name=pept-2; OrderedLocusNames=EF3080;
 OS Enterococcus faecalis (Streptococcus faecalis).
 OC Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.
 OX NCBI_TaxID=1351;
 RN [1]

SEQUENCE FROM N.A.
 RC STRAIN=V583 / ATCC 700802;
 RX MEDLINE=2250857; PubMed=12663927; DOI=10.1126/science.1080613;
 RA Paulsen I.T., Banerjee L., Myers G.S.A., Nelson K.E., Seshadri R.,
 RA Read T.D., Fouts D.E., Eisen J.A., Gill S.R., Heidelberg J.F.,
 RA Tettelin H., Dodson R.J., Umayam L.A., Brinkac L.M., Beanan M.J.,
 RA Daugherty S.C., DeBoy R.T., Durkin S.A., Kolonay J.F., Madupu R.,
 RA Nelson W.C., Vamathevan J.J., Tran B., Upton J., Hansen T., Shetty J.,
 RA Khouri H.M., Utterback T.R., Radune D., Ketchum K.A., Dougherty B.A.,
 RA Fraser C.M.;
 RT "Role of mobile DNA in the evolution of vancomycin-resistant
 RT Enterococcus faecalis";
 RL Science 299:2071-2074 (2003).
 DR EMBL; AE016956; AAC02761.1; -;
 DR HSSP; P29745; IVIX.
 DR TIGR; EF3080; -;
 DR GO; GO:0008237; F:metallopeptidase activity; IEA.
 DR GO; GO:0005508; P:proteolysis and peptidolysis; IEA.
 DR InterPro; IPR001261; Arge_dapE.
 DR InterPro; IPR010161; Peptidase_M20B.
 DR TIGRFAMs; TIGR01882; peptidase-T; 1.
 DR PROSITE; PS00758; ARGE_DAPE_CPG2_1; 1.
 DR PROSITE; PS00759; ARGE_DAPE_CPG2_2; 1.
 KW Complete proteome.
 SQ SEQUENCE 406 AA; 44953 MW; 038ABFCD17BA85DE CRC64;

Query Match 86.2%; Score 25; DB 2; Length 406;
 Best Local Similarity 71.4%; Pred. No. 3.8e+02;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
 : : : : :
 Db 111 VLSVAEF 117

RESULT 20
 Q7VMA4 PRELIMINARY; PRT; 617 AA.
 ID Q7VMA4; DSM 3638 / ATCC 43587 / JCM 8422;
 AC Q7VMA4; Weiss R.B., Dunn D.M., Robb F.T., Brown J.R.;
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Chaperrone protein Hsca.
 GN Name=hscA; OrderedLocusNames=HDI087;
 OS Haemophilus ducreyi.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
 OC Pasteurellaceae; Haemophilus.
 OX NCBI_TaxID=730;
 RN [1]

SEQUENCE FROM N.A.
 RC STRAIN=35000HP / ATCC 700724;
 RA Munson R.S. Jr., Ray W.C., Mahairas G., Sabo P., Mungur R.,
 RA Johnson L., Nguyen D., Wang J., Forst C., Hood L.;
 RT "The complete genome sequence of Haemophilus ducreyi."
 RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: Belongs to the heat shock protein 70 family.
 DR EMBL; AE017153; AAP95953.1; -;
 DR HSSP; P04475; IDKG.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0016887; F:ATPase activity; IEA.
 DR GO; GO:0051082; F:unfolded protein binding; IEA.
 DR GO; GO:0016226; P:iron-sulfur cluster assembly; IEA.
 DR InterPro; IPR010236; Hsca.
 DR InterPro; IPR001023; Hsp70.
 DR Pfam; PF00012; HSP70; 1.
 DR ProDom; PD000089; Hsp70; 1.
 DR TIGRFAMs; TIGR01991; Hsca; 1.
 DR PROSITE; PS00297; HSP70_1; 1.
 DR PROSITE; PS00329; HSP70_2; 1.
 DR PROSITE; PS01036; HSP70_3; 1.
 KW ATP-binding; Complete proteome.
 SQ SEQUENCE 617 AA; 66569 MW; 20512A41BAC1E599 CRC64;

```
Query Match      86.2%; Score 25; DB 2; Length 617;
Best Local Similarity 71.4%; Pred. No. 5.7e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 VMXVAEF 7
DB      553 VLSVAEF 559

RESULT 21
ETFD SCHPO      STANDARD;      PRT;      632 AA.
AC      P8711;
DT      01-NOV-1997 (Rel. 35, Created)
DT      01-NOV-1997 (Rel. 35, Last sequence update)
DT      25-OCT-2004 (Rel. 45, Last annotation update)
DE      Probable electron transfer flavoprotein-ubiquinone oxidoreductase,
DE      mitochondrial precursor (EC 1.5.5.1) (ETP-QO) (ETP-ubiquinone
DE      oxidoreductase) (ETP dehydrogenase) (Electron-transferring-
DE      flavoprotein dehydrogenase).
GN      ORFNames=SPAC20G8.04C;
OS      Schizosaccharomyces pombe (Fission yeast).
OC      Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC      Schizosaccharomycetales; Schizosaccharomycetaceae;
OC      Schizosaccharomycetes.
OX      NCBI_TaxID=4896;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=972;
RX      MEDLINE=21848401; PubMed=11859360; DOI=10.1038/nature724;
RA      Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA      Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA      Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA      Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA      Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA      Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA      James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA      Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odeil C.,
RA      Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA      Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA      Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA      Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA      Woodward J., Volckaert G., Art R., Robben J., Gymonprez B.,
RA      Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA      Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA      Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA      Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA      Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA      Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA      Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA      Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA      Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Foreburg S.L.,
RA      Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA      Spatovski G.V., Usery D., Barrell B.G., Nurse P.;
RT      "The genome sequence of Schizosaccharomyces pombe.";
RL      Nature 415:871-880(2002).
CC      -1- FUNCTION: Accepts electrons from ETP and reduces ubiquinone (By
CC      similarity).
CC      -1- CATALYTIC ACTIVITY: Reduced electron-transferring flavoprotein +
CC      ubiquinone = electron-transferring flavoprotein + ubiquinol.
CC      -1- COFACTOR: FAD and a 4Fe-4S cluster (By similarity).
CC      -1- SUBCELLULAR LOCATION: Mitochondrial inner membrane (By
CC      similarity).
CC      -1- SIMILARITY: Belongs to the ETP-QO / fixC family.
-----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (see http://www.ebi.ac.uk/announcements/
CC      or send an email to license@ebi.ac.uk).
```

```
DR      EMBL; Z95334; CAB08598.1; -.
DR      PIR; T38126; T38126.
DR      GeneDB SPombe; SPAC20G8.04C; -.
DR      InterPro; IPR007859; ETFD.
DR      Pfam; PF05187; ETF_QO; 1.
KW      4Fe-4S; Electron transport; FAD; Flavoprotein; Iron-sulfur;
KW      Mitochondrion; Oxidoreductase; Transit peptide; Ubiquinone.
FT      TRANSIT 1 ? Mitochondrion (Potential).
FT      CHAIN 1 ? Probable electron transfer flavoprotein-
FT      ubiquinone oxidoreductase.
FT      NP BIND 93 107 FAD (ADP part) (Potential).
FT      METAL 575 575 Iron-sulfur (4Fe-4S) (Potential).
FT      METAL 601 601 Iron-sulfur (4Fe-4S) (Potential).
FT      METAL 604 604 Iron-sulfur (4Fe-4S) (Potential).
FT      METAL 607 607 Iron-sulfur (4Fe-4S) (Potential).
SQ      SEQUENCE 632 AA; 69472 MW; 1B0F22374E33771B CRC64;

Query Match      86.2%; Score 25; DB 1; Length 632;
Best Local Similarity 71.4%; Pred. No. 5.8e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 VMXVAEF 7
DB      193 VMSLAEF 199

RESULT 22
Q7UYJ7
ID      Q7UYJ7 PRELIMINARY; PRT; 633 AA.
AC      Q7UYJ7;
DT      01-OCT-2003 (TrEMBLrel. 25, Created)
DT      01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT      01-WAR-2004 (TrEMBLrel. 26, Last annotation update)
DE      1,4-alpha-glucan branching enzyme (EC 2.4.1.18).
GN      Name=glbB; OrderedLocustNames=RB548;
OS      Rhodospirillum rubrum.
OC      Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales;
OC      Planctomycetaceae; Pirellula.
OX      NCBI_TaxID=117;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=1;
RX      MEDLINE=22735913; PubMed=12835416; DOI=10.1073/pnas.1431443100;
RA      Gloeckner F.O., Kube M., Bauer M., Teeling H., Lombardot T.,
RA      Ludwig W., Gade D., Beck A., Borzym K., Heitmann K., Rabus R.,
RA      Schlesner H., Amann R., Reinhardt R.;
RT      "Complete genome sequence of the marine planctomycete Pirellula sp.
RT      strain 1.";
RL      Proc. Natl. Acad. Sci. U.S.A. 100:8298-8303(2003).
DR      EMBL; BX294133; CAD71645.1; -.
DR      GO; GO:0003844; F.1.4-alpha-glucan branching enzyme activity; IEA.
DR      GO; GO:0004556; F.1.4-alpha-glucan branching enzyme activity; IEA.
DR      GO; GO:0004553; F.1.4-alpha-glucan branching enzyme activity; IEA.
DR      GO; GO:0016757; F.1.4-alpha-glucan branching enzyme activity; IEA.
DR      GO; GO:0005975; P.1.4-alpha-glucan branching enzyme activity; IEA.
DR      InterPro; IPR006047; Alpha_amyl_cat.
DR      InterPro; IPR004193; Glyco_hydro_13N.
DR      InterPro; IPR010916; TONB_Box_N.
DR      Pfam; PF00128; Alpha-amylase; 1.
DR      Pfam; PF02922; Isomylase N; 1.
DR      PROSITE; PS00430; TONB_DEPENDENT_REC_1; UNKNOWN 1.
KW      Complete proteome; Glycosyltransferase; Transferase.
SQ      SEQUENCE 633 AA; 72019 MW; 708ECC306CC9D9E7 CRC64;

Query Match      86.2%; Score 25; DB 2; Length 633;
Best Local Similarity 71.4%; Pred. No. 5.8e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 VMXVAEF 7
DB      198 IMPVAEF 204
```

```
RESULT 23
P92974 PRELIMINARY; PRT; 1077 AA.
AC P92974;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 03-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Ubiquitin activating enzyme 2.
GN Name=UBA2;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RN SEQUENCE FROM N.A.
RP MEDLINE=98069011; PubMed=9405937;
RA Kotani H., Nakamura Y., Sato S., Kaneko T., Asamizu E., Miyajima N.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. II.
RT Sequence features of the regions of 1,044,062 bp covered by thirteen
RT physically assigned P1 clones.";
RL DNA Res. 4:291-300(1997).
DR ENBL; U40566; AAB37569.1; -.
DR ENBL; AB006700; BAB08968.1; -.
DR HSSP; P12282; 1JW9.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0004839; F:ubiquitin activating enzyme activity; IEA.
DR GO; GO:0006512; P:ubiquitin cycle; IEA.
DR InterPro; IPR008262; Lipase_AS.
DR InterPro; IPR009036; Moeb.
DR InterPro; IPR002020; NAD_BS.
DR InterPro; IPR000594; Thif domain.
DR InterPro; IPR000127; UBact repeat.
DR InterPro; IPR000011; Ugtin-activ_enz.
DR Pfam; PF00899; Thif; 2.
DR Pfam; PF02134; UBACT; 2.
DR PRINTS; PR01849; UBIQUITINACT.
DR TIGRFAMS; TIGR01408; Ubel; 1.
DR PROSITE; PS00120; LIPASE_SER; UNKNOWN 1.
DR PROSITE; PS00536; UBIQUITIN_ACTIVAT_1; 1.
DR PROSITE; PS00865; UBIQUITIN_ACTIVAT_2; 1.
SQ SEQUENCE 1077 AA; 119622 MW; CE39A36AA99A218 CRC64;

Query Match 86.2%; Score 25; DB 2; Length 1077;
Best Local Similarity 71.4%; Pred. No. 9.6e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
Db 54 VMSMAEF 60

RESULT 24
Q6Y252 PRELIMINARY; PRT; 140 AA.
AC Q6Y252;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 03-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE ATPase class II (Fragment).
OS Pagrus major (Red sea bream) (Chrysophrys major).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Anthomorpha; Acanthopterygii; Perciformes; Percoidae;
OC Sparidae; Pagrus.
OX NCBI_TaxID=143350;
RN [1]
```

```
RP SEQUENCE FROM N.A.
RC TISSUE=Spleen;
RA Chen S.L., Xu M.Y., Hu S.L., Li L.;
RT "Analysis of immune-relevant genes expressed in red sea bream
RT spleen.";
RL Aquaculture 240:115-130(2004).
DR EMBL; AY190684; AAP20160.1; -.
FT NON TER 1
SQ SEQUENCE 140 AA; 15723 MW; 20EFF0F49C40EB09 CRC64;

Query Match 82.8%; Score 24; DB 2; Length 140;
Best Local Similarity 71.4%; Pred. No. 2.7e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
Db 72 LMTVAEF 78

RESULT 25
RL13 HALMA STANDARD; PRT; 145 AA.
AC P29138;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE 50S ribosomal protein L13p (Hmal13).
GN Name=rp13p;
OS Haloarcula marismortui (Halo bacterium marismortui).
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC Halobacteriaceae; Haloarcula.
OX NCBI_TaxID=2238;
RN [1]
RN SEQUENCE FROM N.A.
RP MEDLINE=92105119; PubMed=1840597;
RA Kroemer W.J., Arndt E.;
RT "Halo bacterial S9 operon. Three ribosomal protein genes are
RT cotranscribed with genes encoding a tRNA(Leu), the enolase, and a
RT putative membrane protein in the archaeobacterium Haloarcula
RT (Halo bacterium) marismortui";
RL J. Biol. Chem. 266:24573-24579(1991).
RN [2]
RN X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF THE 50S SUBUNIT.
RP STRAIN=ATCC 43049;
RX MEDLINE=20396344; PubMed=10937989; DOI=10.1126/science.289.5481.905;
RA Ban N., Nissen P., Hansen J., Moore P.B., Steitz T.A.;
RT "The complete atomic structure of the large ribosomal subunit at 2.4 A
RT resolution.";
RL Science 289:905-920(2000).
RN [3]
RN X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS) OF THE 50S SUBUNIT.
RP STRAIN=ATCC 43049;
RX PubMed=10937990; DOI=10.1126/science.289.5481.920;
RA Nissen P., Hansen J., Ban N., Moore P.B., Steitz T.A.;
RT "The structural basis of ribosome activity in peptide bond
RT synthesis.";
RL Science 289:920-930(2000).
RN [4]
RN X-RAY CRYSTALLOGRAPHY (3.1 ANGSTROMS) OF THE 50S SUBUNIT.
RP STRAIN=ATCC 43049;
RX PubMed=11828326; DOI=10.1038/nsb758;
RA Schmeing T.M., Seila A.C., Hansen J.L., Freeborn B., Soukup J.K.,
RA Schmeing S.A., Strobel S.A., Moore P.B., Steitz T.A.;
RT "A pre-translational intermediate in protein synthesis observed in
RT crystals of enzymatically active 50S subunits.";
RL Nat. Struct. Biol. 9:225-230(2002).
RN [5]
RN X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF THE 50S SUBUNIT.
RP STRAIN=ATCC 43049;
RX PubMed=11483524; DOI=10.1093/emboj/20.15.4214;
RA Klein D.J., Schmeing T.M., Moore P.B., Steitz T.A.;
RT "The kink-turn: a new RNA secondary structure motif.";
RL EMBO J. 20:4214-4221(2001).
```

[6]
 RP X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS) OF THE 50S SUBUNIT IN COMPLEX
 RP WITH FOUR MACROLIDE ANTIBIOTICS.
 RC STRAIN=ATCC 43049;
 RX PubMed=12150912; DOI=10.1016/S1097-2765(02)00570-1;
 RA Hansen J.L., Ippolito J.A., Ban N., Nissen P., Moore P.B.,
 RA Steitz T.A.;
 RT "The structures of four macrolide antibiotics bound to the large
 RT ribosomal subunit."; <http://www.sciencedirect.com/science/article/pii/S0969212602000128>
 RL Mol. Cell 10:117-128 (2002).
 [7]
 RN X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF THE 50S SUBUNIT.
 RC STRAIN=ATCC 43049;
 RX PubMed=12185246; DOI=10.1073/pnas.172404099;
 RA Hansen J.L., Schmeing T.M., Moore P.B., Steitz T.A.;
 RT "Structural insights into peptide bond formation."; <http://www.pnas.org/cgi/doi/10.1073/pnas.172404099>
 RL Proc. Natl. Acad. Sci. U.S.A. 99:11670-11675 (2002).
 [8]
 RN X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS) OF THE 50S SUBUNIT IN COMPLEX
 RP WITH FIVE ANTIBIOTICS AT THE PEPTIDYL TRANSFERASE CENTER.
 RC STRAIN=ATCC 43049;
 RX PubMed=12860128; DOI=10.1016/S0022-2836(03)00668-5;
 RA Hansen J.L., Moore P.B., Steitz T.A.;
 RT "Structures of five antibiotics bound at the peptidyl transferase
 RT center of the large ribosomal subunit."; <http://www.sciencedirect.com/science/article/pii/S0022283603006685>
 RL J. Mol. Biol. 330:1061-1075 (2003).
 [9]
 RN X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS) OF THE 50S SUBUNIT WITH TWO
 RP DIFFERENT E SITE SUBSTRATES.
 RX PubMed=14561884; DOI=10.1261/rna.5120503;
 RA Schmeing T.M., Moore P.B., Steitz T.A.;
 RT "Structures of deacylated tRNA mimics bound to the E site of the large
 RT ribosomal subunit."; <http://www.rna-journal.org/cgi/doi/10.1261/rna.5120503>
 RL RNA 9:1345-1352 (2003).
 CC -1- FUNCTION: Binds to the 23S rRNA.
 CC -1- SUBUNIT: Part of the 50S ribosomal subunit. Interacts weakly with
 CC proteins L3 and L6.
 CC -1- SIMILARITY: Belongs to the ribosomal protein L13p family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M76567; AAA73097.1; -;
 DR PIR; B41715; B41715.
 DR PDB; 1FFK; X-ray; G=-.
 DR PDB; 1J02; X-ray; I=1-145.
 DR PDB; 1K73; X-ray; K=-.
 DR PDB; 1K8A; X-ray; K=1-145.
 DR PDB; 1K9M; X-ray; K=1-145.
 DR PDB; 1KC8; X-ray; K=1-145.
 DR PDB; 1KD1; X-ray; K=1-145.
 DR PDB; 1KQ5; X-ray; I=-.
 DR PDB; 1MLK; X-ray; K=1-145.
 DR PDB; 1M90; X-ray; K=1-145.
 DR PDB; 1N8R; X-ray; K=1-145.
 DR PDB; 1NJI; X-ray; K=1-145.
 DR PDB; 1Q71; X-ray; K=1-145.
 DR PDB; 1Q81; X-ray; K=1-145.
 DR PDB; 1Q82; X-ray; K=1-145.
 DR PDB; 1Q86; X-ray; K=1-145.
 DR PDB; 1QVF; X-ray; I=1-145.
 DR PDB; 1QVG; X-ray; I=1-145.
 DR InterPro; IPR005822; Ribosomal_L13.
 DR InterPro; IPR005755; Ribosomal_L13e/a.
 DR Pfam; PF00572; Ribosomal_L13; 1.
 DR ProDom; PD001791; Ribosomal_L13; 1.
 DR TIGRfams; TIGR01077; L13 A.E; 1.
 DR PROSITE; PS00783; RIBOSOMAL_L13; 1.

KW 3D-structure; Ribosomal protein; RNA-binding; rRNA-binding.
 SQ SEQUENCE 145 AA; 16228 MW; 069CE66662AE3BC CRC64;
 Query Match 82.8%; Score 24; DB 1; Length 145;
 Best Local Similarity 83.3%; Pred. No. 2.7e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 2 MKVAEF 7
 DB 1 MSVAEF 6
 RESULT 26
 Q72IL7 PRELIMINARY; PRT; 148 AA.
 AC Q72IL7;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Hypothetical protein.
 GN OrderedLocNames=TTCl115;
 OS Thermus thermophilus (strain HB27 / AFCC BAA-163 / DSM 7039) .
 OC Bacteria; Deinococcus-Thermus; Deinococci; Thermales; Thermaceae;
 OC Thermus.
 OX NCBI_TaxID=262724;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX PubMed=15064768;
 RA Henne A., Brueggemann H., Raasch C., Wierse A., Hartsch T.,
 RA Liesegang H., Johann A., Lienard T., Gohl O., Martinez-Arias R.,
 RA Jacobi C., Starkuviene V., Schlenszcek S., Dencker S., Huber R.,
 RA Klenk H.-P., Kramer W., Merkl R., Gottschalk G., Fritz H.-J.;
 RT "The genome sequence of the extreme thermophile Thermus
 RT thermophilus."; <http://www.ncbi.nlm.nih.gov/pubmed/15064768>
 RL Nat. Biotechnol. 22:547-553 (2004) .
 DR EMBL; AA017304; AAS81457.1; -;
 DR InterPro; IPR009072; Histone-fold.
 KW Complete proteome; Hypothetical protein.
 SQ SEQUENCE 148 AA; 16860 MW; 0A540838B46F96AB CRC64;
 Query Match 82.8%; Score 24; DB 2; Length 148;
 Best Local Similarity 71.4%; Pred. No. 2.8e+02;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 VMKVAEF 7
 DB 2 LMKVAEF 8
 RESULT 27
 Q8TLY8 PRELIMINARY; PRT; 149 AA.
 AC Q8TLY8;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Hypothetical protein MA2884.
 GN OrderedLocNames=MA2884;
 OS Methanosarcina acetivorans.
 OC Archaea; Euryarchaeota; Methanomicrobia; Methanosarcinales;
 OC Methanosarcinaceae; Methanosarcina.
 OX NCBI_TaxID=2214;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C2A / ATCC 35395 / DSM 2834;
 RX MEDLINE=21929760; PubMed=11932238; DOI=10.1101/gr.223902;
 RA Galagan J.E., Nusbaum C., Roy A., Endrizzi M.G., Macdonald P.,
 RA FitzHugh W., Calvo S., Engels R., Smirnov S., Anoor D., Brown A.,
 RA Allen N., Naylor J., Stange-Thomann N., DeArellano K., Johnson R.,
 RA Linton L., McEwan P., McKernan K., Talamas J., Tirrell A., Ye W.,
 RA Zimmer A., Barber R.D., Cann I., Graham D.E., Grubbs D.A., Guss A.M.,
 RA Hedderich R., Ingram-Smith C., Kuettnner H.C., Krzycki J.A.,
 RA Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,


```

RA Springer T.A., Umayam L.A., White O., White R.H., de Macario E.C.,
RA Ferry J.G., Jarrell K.F., Jing H., Macario A.J.L., Paulsen I.T.,
RA Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,
RA Metcalf W.W., Birren B.;
RT "The genome of Methanococcus aceticivorans reveals extensive metabolic
RT and physiological diversity.";
RL Genome Res. 12:532-542(2002).
DR EMBL; AE010989; ANM06261.1; -.
DR GO; GO:0008168; F:methyltransferase activity; IEA.
DR InterPro; IPR001601; Methyltransf.
KW Complete proteome.
SQ SEQUENCE 149 AA; 16402 MW; EF8576A410EC13F3 CRC64;

Query Match      82.8%; Score 24; DB 2; Length 149;
Best Local Similarity 71.4%; Pred. No. 2.9e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
DB 13 VLGVAEF 19

RESULT 28
O86ZD3 PRELIMINARY; PRT; 154 AA.
AC Q86ZD3;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative succinyl-CoA synthase (Fragment).
GN Name=Fspscs;
OS Fusarium sp. IFO 7772.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Hypocreomycetidae; Hypocreales; mitosporic Hypocreales; Fusarium.
OX NCBI_TaxID=182099;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IFO 7772;
RX MEDLINE=22505401; PubMed=12618405;
RA Kimura M., Tokai T., Matsumoto G., Fujimura M., Hamamoto H.,
RA Yoneyama K., Shibata T., Yamaguchi I.;
RT "Trichothecene nonproducer Gibberella species have both functional and
RT nonfunctional 3-O-acetyltransferase genes.";
RL Genetics 163:677-684(2003).
DR EMBL; AB076255; BAC65219.1; -.
DR HSSP; P07459; 2SCU
DR InterPro; IPR003781; CoA_binding.
DR Pfam; PF02629; CoA_binding; 1.
FT NON_TER 154
FT TER 154
SQ SEQUENCE 154 AA; 16632 MW; F46D6A325AD1EEC1 CRC64;

Query Match      82.8%; Score 24; DB 2; Length 154;
Best Local Similarity 71.4%; Pred. No. 2.9e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
DB 124 VVSVAEF 130

RESULT 29
ID _RAP_TAROF STANDARD; PRT; 157 AA.
AC O49065;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 29-MAR-2004 (Rel. 43, Last annotation update)
DE Root allergen protein (RAP).
OS Taraxacum officinale (Common dandelion).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC campanulids; Asterales; Asteraceae; Cichorioideae; Cichorieae;
OC Taraxacum.

```

```

OX NCBI_TaxID=50225;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Root;
RA Xu X.-Y., Bewley J.D., Greenwood J.S.;
RL Submitted (DEC1997) to the EMBL/GenBank/DBJ databases.
CC -!- ALLERGEN: Causes an allergic reaction in human.
CC -!- SIMILARITY: Belongs to the BetV1 family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF036931; AAB92255.1; -.
CC HSSP; O24248; 1E09.
CC InterPro; IPR000916; Bet_v_1.
CC Pfam; PF00407; Bet_v_1; 1.
CC PRINTS; PR00634; BETALLERGEN.
CC ProDom; PD000531; Bet_v_1; 1.
CC PROSITE; PS00451; PATHOGENESIS_BETV1; 1.
KW Allergen; Pathogenesis-related protein; Plant defense.
SQ SEQUENCE 157 AA; 17040 MW; 5892AB8593A8A7E0 CRC64;

Query Match      82.8%; Score 24; DB 1; Length 157;
Best Local Similarity 83.3%; Pred. No. 3e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 MXVAEF 7
DB 1 MAVAEF 6

RESULT 30
O89G21 PRELIMINARY; PRT; 161 AA.
AC O89G21;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Bll6527 protein.
GN OrderedLocusNames=bll6527;
OS Bradyrhizobium japonicum.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Bradyrhizobium.
OX NCBI_TaxID=375;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=USDA110;
RX MEDLINE=22484998; PubMed=12597275;
RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiumi T.,
RA Sasamoto S., Watanabe A., Idesawa K., Iriguchi M., Kawashima K.,
RA Kohara M., Matsumoto M., Shimo S., Tsuruoka H., Wada T., Yamada M.,
RA Tabata S.;
RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
RT Bradyrhizobium japonicum USDA110.";
RL DNA Res. 9:189-197(2002).
DR EMBL; AF005958; BAC51792.1; -.
KW Complete proteome.
SQ SEQUENCE 161 AA; 17456 MW; 4B7895C4BF28D199 CRC64;

Query Match      82.8%; Score 24; DB 2; Length 161;
Best Local Similarity 71.4%; Pred. No. 3e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
DB 68 VLGVAEF 74

```

```

RESULT 31
Q6AQI2 PRELIMINARY; PRT; 162 AA.
ID Q6AQI2
AC Q6AQI2
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=DP0662;
OS Desulfotalea psychrophila.
OC Bacteria; Proteobacteria; Deltaproteobacteria; Desulfobacteriales;
OC Desulfobulbaceae; Desulfotalea.
OX NCBI_TaxID=84980;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LSV54 / DSM 12343;
RX PubMed=15305914;
RA Rabus R., Ruepp A., Frickey T., Rattei T., Fartmann B., Stark M.,
RA Bauer M., Zibat A., Lombardot T., Becker I., Amann J., Gellner K.,
RA Teeling H., Leuschner W.D., Gloeckner F.-O., Lupas A.N., Amann R.,
RA Klenk H.-P.;
RT "The genome of Desulfotalea psychrophila, a sulfate-reducing bacterium
RT from permanently cold Arctic sediments.";
RL Environ. Microbiol. 6:887-902 (2004).
DR EMBL: CR522870; CAG35391.1; -.
DR InterPro: IPR009921; DUF1456.
DR Pfam: PF07308; DUF1456; 1.
KW Complete proteome.
SQ SEQUENCE 162 AA; 19034 MW; 5E67041F45744728 CRC64;

Query Match 82.8%; Score 24; DB 2; Length 162;
Best Local Similarity 57.1%; Pred. No. 3e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
Db 114 IMAAEF 120

RESULT 32
Q8SRD1 PRELIMINARY; PRT; 165 AA.
ID Q8SRD1
AC Q8SRD1
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE COATOMER ZETA SUBUNIT.
GN Names=ECU08_0680;
OS Encephalitozoon cuniculi GB-M1.
OC Eukaryota; Fungi; Microsporidia; Unikaryonidae; Encephalitozoon.
OX NCBI_TaxID=284813;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GB-M1;
RX MEDLINE=21576510; PubMed=11719806; DOI=10.1038/35106579;
RA Katinka M.D., Duprat S., Cornillot E., Metenier G., Thonart P.,
RA Prensier G., Barbe V., Peyrecaillade E., Brottier P., Wincker P.,
RA Delbac F., El Alaoui H., Peyret P., Saurin W., Gouy M.,
RA Weissenbach J., Vivares C.P.;
RT "Genome sequence and gene compaction of the eukaryote parasite
RT Encephalitozoon cuniculi.";
RL Nature 414:450-453 (2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=GB-M1;
RA Genoscope;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL590448; CAD26373.1; -.
DR GO: GO:0006810; P:transport; IEA.
DR InterPro: IPR011012; Longin-like.
SQ SEQUENCE 165 AA; 18742 MW; 7CDAB6C35BC11A6C CRC64;

Query Match 82.8%; Score 24; DB 2; Length 165;

Best Local Similarity 71.4%; Pred. No. 3.1e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
Db 107 VLLVAEF 113

RESULT 33
SSB_RHOBA STANDARD; PRT; 169 AA.
ID SSB_RHOBA
AC P59332; Q70KV3;
DT 29-MAR-2004 (Rel. 43, Created)
DT 29-MAR-2004 (Rel. 43, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Single-strand binding protein (SSB) (Helix-destabilizing protein).
GN Name=ssb; OrderedLocusNames=RB9917;
OS Rhodopirellula baltica.
OC Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales;
OC Planctomycetaceae; Pirellula.
OX NCBI_TaxID=117;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1;
RX MEDLINE=22735913; PubMed=12835416; DOI=10.1073/pnas.1431443100;
RA Gloeckner F.O., Kube M., Bauer M., Teeling H., Lombardot T.,
RA Ludwig W., Gade D., Beck A., Borzym K., Heitmann K., Rabus R.,
RA Schlesner H., Amann R., Reinhardt R.;
RT "Complete genome sequence of the marine planctomycete Pirellula sp.
RT strain 1.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:8298-8303 (2003).
CC -!- FUNCTION: This protein is essential for replication of the
CC chromosome. It is also involved in DNA recombination and repair
CC (by similarity).
CC -!- SIMILARITY: Contains 1 SSB domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: BX294150; CAD76529.1; -.
CC InterPro: IPR010913; SS_binding.
CC InterPro: IPR000424; SSB_protein.
CC Pfam: PF04336; SSB; 1.
CC TIGRFAMs: TIGR00621; ssb; 1.
CC PROSITE: PS50935; SSB; 1.
CC Complete proteome; DNA repair; DNA replication; DNA-binding.
FT DOMAIN 4 107 SSB.
FT DOMAIN 113 121 Poly-Gly.
SQ SEQUENCE 169 AA; 18108 MW; AA8F93B0FA51F287 CRC64;

Query Match 82.8%; Score 24; DB 1; Length 169;
Best Local Similarity 83.3%; Pred. No. 3.2e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 MXVAEF 7
Db 26 MAVAEF 31

RESULT 34
Q6M2B7 PRELIMINARY; PRT; 169 AA.
ID Q6M2B7
AC Q6M2B7
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Multisubunit Na+/H+ antiporter, subunit E.
GN Name=mrpE; OrderedLocusNames=cg3027;

```

OS Corynebacterium glutamicum (Brevibacterium flavum).
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
 OX NCBI_TaxID=1718;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
 RX PubMed=12948626; DOI=10.1016/S0168-1656(03)00154-8;
 RA Kalinowski J., Bathe B., Bartels D., Bischoff N., Bott M.,
 RA Burkovski A., Dusch N., Eggeling L., Eikmanns B.J., Gaigalat L.,
 RA Goemann A., Hartmann M., Huthmacher K., Kraemer R., Linke B.,
 RA McHardy A.C., Meyer F., Moechel B., Pfeifferle W., Puhler A.,
 RA Rey D.A., Rueckert C., Rupp O., Sahn H., Wendisch V.F., Wiegand I.,
 RA Tauch A.;
 RT "The complete Corynebacterium glutamicum ATCC 13032 genome sequence
 RT and its impact on the production of L-aspartate-derived amino acids
 RT and vitamins";
 RL J. Biotechnol. 104:5-25 (2003).
 DR EMBL; BX927156; CAF20754.1; -.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0008324; P:cation transporter activity; IEA.
 DR GO; GO:0006812; P:cation transport; IEA.
 DR InterPro; IPR002758; Cation_antiport.
 DR ProDom; PD012569; DUF68; 1.
 SQ SEQUENCE 169 AA; 19166 MW; CB9AA69662E22A36 CRC64;
 Query Match 82.8%; Score 24; DB 2; Length 169;
 Best Local Similarity 71.4%; Pred. No. 3.2e+02;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 VMXVAEF 7
 Db 22 VMLMAEF 28

RESULT 35
 Q9N416
 ID Q9N416 PRELIMINARY; PRT; 189 AA.
 AC Q9N416;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Hypothetical protein Y71F9AL.10.
 GN Names=Y71F9AL.10; ORFNames=Y71F9AL.10;
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RX MEDLINE=99069613; PubMed=9851916;
 RG WormBase Consortium;
 RT "Genome sequence of the nematode C. elegans: a platform for
 RT investigating biology. The C. elegans Sequencing Consortium.";
 RL Science 282:2012-2018 (1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2.
 RA Bradshaw-Cordum H., Scott K., Graves T.;
 RT "The sequence of C. elegans cosmid Y71F9AL.";
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Waterston R.;
 RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Wilson R.;
 RL Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.
 RN [5]

RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Wilson R.;
 RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RG WormBase Consortium;
 RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: Contains 1 RING-type zinc finger.
 DR EMBL; AC024200; AAF35997.2; -.
 DR HSP; Q9LRB7; Ilym.
 DR WormBase; WBGene0022115; Y71F9AL.10.
 DR WormPep; Y71F9AL.10; CE31386.
 DR GO; GO:0001511; C:ubiquitin ligase complex; IEA.
 DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.
 DR GO; GO:0008270; F:zinc ion binding; IEA.
 DR GO; GO:0016567; P:protein ubiquitination; IEA.
 DR InterPro; IPR001841; Znf_ring.
 DR Pfam; PF00097; zf-C3HC4; 1.
 DR SMART; SM00184; RING; 1.
 DR PROSITE; PS00089; ZF_RING_2; 1.
 KW Hypothetical protein; Metal-binding; Zinc; Zinc-finger.
 SQ SEQUENCE 189 AA; 21048 MW; 7BDA2D05F0362CD0 CRC64;
 Query Match 82.8%; Score 24; DB 2; Length 189;
 Best Local Similarity 71.4%; Pred. No. 3.5e+02;
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 VMXVAEF 7
 Db 37 VMTAAEF 43

RESULT 36
 Q97F65
 ID Q97F65 PRELIMINARY; PRT; 195 AA.
 AC Q97F65;
 DT 01-OCT-2001 (TrEMBLrel. 18, Created)
 DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Thymidine kinase.
 GN OrderedLocustNames=CAC2887;
 OS Clostridium acetobutylicum.
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
 OC Clostridium.
 OX NCBI_TaxID=1488;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
 RX MEDLINE=21359325; PubMed=11466286;
 RX DOI=10.1128/JB.183.16.4823-4838.2001;
 RA Noelling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q.,
 RA Gibson R., Lee H.M., Dubois J., Qiu D., Hitti J., Wolf Y.I.,
 RA Tatusev R.L., Sabathe F., Doucette-Stamm L.A., Soucaille P.,
 RA Daly M.J., Bennett G.N., Koonin E.V., Smith D.R.;
 RT "Genome sequence and comparative analysis of the solvent-producing
 RT bacterium Clostridium acetobutylicum";
 RL J. Bacteriol. 183:4823-4838 (2001).
 CC -1- CATALYTIC ACTIVITY: ATP + thymidine = ADP + thymidine 5'-
 CC phosphate.
 CC -1- SIMILARITY: Belongs to the thymidine kinase family.
 DR EMBL; AE007786; AAK80830.1; -.
 DR PIR; C97255; C97255.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0016301; F:kinase activity; IEA.
 DR GO; GO:0004797; F:thymidine kinase activity; IEA.
 DR GO; GO:0016740; F:transferase activity; IEA.
 DR GO; GO:0006259; P:DNA metabolism; IEA.
 DR InterPro; IPR00345; CytC_heme_BS.
 DR InterPro; IPR001267; TK_cell.
 DR Pfam; PF00265; TK; 1.
 DR PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_1.

```

DR PROSITE: PS00603; TK CELLULAR TYPE; UNKNOWN 1.
KW ATP-binding; Complete proteome; DNA synthesis; Kinase; Transferase.
SQ SEQUENCE 195 AA; 22303 MW; 6ED551986B203BA CRC64;

Query Match      82.8%; Score 24; DB 2; Length 195;
Best Local Similarity 57.1%; Pred. No. 3.6e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEP 7
Db 131 LMAIAEP 137

RESULT 37
Q8XWK7 PRELIMINARY; PRT; 206 AA.
AC Q8XWK7
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE PROBABLE TRANSMEMBRANE PROTEIN.
GN Name=RS01138; OrderedLocusNames=RS02467;
OS Ralstonia solanacearum (Pseudomonas solanacearum).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Ralstonia.
OX NCBI_TaxID=305;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GMI1000;
RX MEDLINE=2168179; PubMed=11823852; DOI=10.1038/415497a;
RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
RA Ariat M., Billault A., Brottier P., Camus J.C., Cattolico L.,
RA Chandler M., Choise N., Claudel-Renard C., Cunnac S., Demange N.,
RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schiex T.,
RA Siquier P., Thebaud P., Whalen M., Wincker P., Levy M.,
RA Weissenbach J., Boucher C.A.;
RT "Genome sequence of the plant pathogen Ralstonia solanacearum.";
RL Nature 415:497-502(2002).
DR EMBL; AL646070; CAD16174.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
KW Complete proteome; Transmembrane.
SQ SEQUENCE 206 AA; 21783 MW; 8D49E45A150B8440 CRC64;

Query Match      82.8%; Score 24; DB 2; Length 206;
Best Local Similarity 57.1%; Pred. No. 3.8e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEP 7
Db 55 VLIAEP 61

RESULT 38
YP58_MYCTU STANDARD; PRT; 236 AA.
AC P65005; Q50740;
DT 01-NOV-1997 (Rel. 35, Created)
DT 25-OCT-2004 (Rel. 45, Last sequence update)
DT 25-JAN-2005 (Rel. 46, Last annotation update)
DE Hypothetical protein RV2558/MT2635.
GN OrderedLocusNames=RV2558, MT2635; ORFNames=MYCY9C4.10c;
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RX MEDLINE=98295987; PubMed=9634230; DOI=10.1038/311159;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C.M.,
RA Harris E.E., Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III,
RA Tekala F., Badcock K., Basham D., Brown D., Chillingworth T.,
RA Connor R., Davies R.M., Devlin K., Feltwell T., Gentles S., Hamlin N.,
RA Holroyd S., Hornsby T., Jagels K., Krogh A., McLean J., Moule S.,
RA Murphy L.D., Oliver S., Osborne J., Quail M.A., Rajandream M.A.,
RA Rogers J., Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrett B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
complete genome sequence.";
RL Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RX MEDLINE=22206494; PubMed=12218036;
RX DOI=10.1128/JB.184.19.5479-5490.2002;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J.D., DeBoy R.T., Dodson R.J., Gwinn M.L., Haft D.H.,
RA Hickey B.K., Kolonay J.F., Nelson W.C., Umayam L.A., Emdolaeva M.D.,
RA Salzberg S.L., Delcher A., Utterback T.R., Weidman J.P., Khouri H.M.,
RA Gill J., Mikula A., Bishai W., Jacobs W.R. Jr., Venter J.C.,
RA Fraser C.M.;
RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and
laboratory strains.";
RL J. Bacteriol. 184:5479-5490(2002).
CC -!- SIMILARITY: STRONG, TO M.TUBERCULOSIS RV2557.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; BX842580; CAB01046.1; -.
CC EMBL; AB000516; AAK46947.1; -.
CC PIR; B70728; B70728.
CC TIGR; MT2635; -.
CC TubercuList; RV2558; -.
CC InterPro: IPR011008; Dimer A B barrel.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 236 AA; 25718 MW; 13E3B049D8F79C6B CRC64;

Query Match      82.8%; Score 24; DB 1; Length 236;
Best Local Similarity 71.4%; Pred. No. 4.3e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEP 7
Db 217 VLDVAEP 223

RESULT 39
YP88_MYCBO STANDARD; PRT; 236 AA.
AC P65006; Q50740;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Hypothetical protein MD2588.
GN OrderedLocusNames=Mb2588;
OS Mycobacterium bovis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1765;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AF2122/97;
RX MEDLINE=22709107; PubMed=12788972; DOI=10.1073/pnas.1130426100;
RA Garnier T., Eiglmeier K., Camus J.-C., Medina N., Mansoor H.,
RA Pryor M., Duthoy S., Grondin S., Lacroix C., Monsemp C., Simon S.,
RA Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
RA Parkhill J., Barrett B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
RT "The complete genome sequence of Mycobacterium bovis.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).
CC -!- SIMILARITY: STRONG, TO M.TUBERCULOSIS RV2557.

```

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; BX248343; CAD94773.1; -
DR InterPro; IPR011008; Dimer A.B barrel.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 236 AA; 25718 MW; 13E3B049D8F79C6B CRC64;

Query Match      82.8%; Score 24; DB 1; Length 236;
Best Local Similarity 71.4%; Pred. No. 4.3e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
Db 217 VLDVAEF 223
[ : ||| ]

RESULT 40
Q6NIE1 PRELIMINARY; PRT; 238 AA.
AC Q6NIE1
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=DIP0836;
OS Corynebacterium diphtheriae.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=1717;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Biotype gravis / NCTC 13129;
RX MEDLINE=22965443; PubMed=14602910; DOI=10.1093/nar/sgk874;
RA Cerdeno-Tarraga A.-M., Efstratiou A., Dover L.G., Holden M.T.G.,
RA Pallen M.J., Bentley S.D., Beara G.S., Churcher C.M., James K.D.,
RA De Zoysa A., Chillingworth T., Cronin A., Dowd L., Feltwell T.,
RA Hamlin N., Holroyd S., Jagels K., Moule S., Quail M.A.,
RA Rabinovitch E., Rutherford K.M., Thomson N.R., Unwin L.,
RA Whitehead S., Barrell B.G., Parkhill J.,
RT "The complete genome sequence and analysis of Corynebacterium
RT diphtheriae NCTC13129."
RL Nucleic Acids Res. 31:6516-6523(2003).
DR EMBL; BX248356; CAE49352.1; -
DR GO; GO:0004222; F:metallopeptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR02886; Peptidase_M23B.
DR InterPro; IPR011054; R:hydrolase M23.
DR Pfam; PF01551; Peptidase M23; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 238 AA; 24945 MW; 1560B577A2194F0 CRC64;

Query Match      82.8%; Score 24; DB 2; Length 238;
Best Local Similarity 57.1%; Pred. No. 4.4e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
Db 72 VLSIAEF 78
[ : ||| ]

RESULT 41
O97008 PRELIMINARY; PRT; 248 AA.
ID O97008
AC O97008
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)

```

```

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein L7610.05.
GN Names=L7610.05;
OS Leishmania major.
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5664;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Friedlin;
RX MEDLINE=98146435; PubMed=9477341;
RA Ivens A.C., Lewis S.M., Bagherzadeh A., Zhang L., Chan H.M.,
RA Smith D.F.;
RT "A physical map of the Leishmania major Friedlin genome."
RL Genome Res. 8:135-145(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Friedlin;
RA Ivens A.C., Murphy L., Quail M., Harris D., Oliver K., Lawson D.,
RA Rajandream M.A., Barrell B.G.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL139794; CAC22635.1; -
DR PIR; T18315; T18315.
KW Hypothetical protein.
SQ SEQUENCE 248 AA; 26365 MW; A80440E18B3F69B5 CRC64;

Query Match      82.8%; Score 24; DB 2; Length 248;
Best Local Similarity 71.4%; Pred. No. 4.6e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
Db 196 LMRVAEF 202
[ : ||| ]

RESULT 42
Q9PKD5 PRELIMINARY; PRT; 249 AA.
ID Q9PKD5
AC Q9PKD5
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Serine/threonine protein phosphatase, putative.
GN OrderedLocusNames=TC0530;
OS Chlamydia muridarum.
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83560;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MoPn / Ni99;
RX MEDLINE=20150255; PubMed=10684935; DOI=10.1093/nar/28.6.1397;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA White O., Hickey E.K., Peterson J.D., Utterback T.R., Berry K.J.,
RA Bass S., Linher K.D., Weidman J.F., Khouri H.M., Craven B., Bowman C.,
RA Dodson R.J., Gwinn M.L., Nelson W.C., DeBoy R.T., Kolonay J.F.,
RA McClarty G., Salzberg S.L., Eisen J.A., Fraser C.M.;
RT "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
RT pneumoniae AR39."
RL Nucleic Acids Res. 28:1397-1406(2000).
DR EMBL; AE002321; AAF39372.1; -
DR PIR; G81693; G81693.
DR TIGR; TC0530; -
DR GO; GO:0003824; F:catalytic activity; IEA.
DR InterPro; IPR001932; PF2C-like.
DR Pfam; PF00481; PP2C; 1.
DR SMART; SM00332; PP2C; 1.
DR SMART; SM00331; PP2C_SIG; 1.
KW Complete proteome.
SQ SEQUENCE 249 AA; 28192 MW; 6BD2278D3281164F CRC64;

Query Match      82.8%; Score 24; DB 2; Length 249;
Best Local Similarity 71.4%; Pred. No. 4.6e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 VMXVAEF 7
   : |||||
Db 1 MMIVAEF 7

RESULT 43
Q9NE74 PRELIMINARY; PRT; 270 AA.
AC Q9NE74;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein L5075.02.
GN Name=L5075.02;
OS Leishmania major.
OC Eukaryota; Euzlenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5664;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Friedlin;
RX MEDLINE=98146435; PubMed=9477341;
RA Ivens A.C., Lewis S.M., Bagherzadeh A., Zhang L., Chan H.M.,
RA Smith D.F.;
RT "A physical map of the Leishmania major Friedlin genome.";
RL Genome Res. 8:135-145(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Friedlin;
RA Beck A., Klages S., Reinhardt R., Ivens A.C., Quail M.,
RA Rajandream M.A., Barrell B.G.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ163552; CAB86964.1; -.
KW Hypothetical protein.
SQ SEQUENCE 270 AA; 29335 MW; 5B91DA37A16FC3CB CRC64;

Query Match 82.8%; Score 24; DB 2; Length 270;
Best Local Similarity 71.4%; Pred. No. 4.9e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
   : |||||
Db 148 VMNLAEF 154

RESULT 44
Q661U0 PRELIMINARY; PRT; 297 AA.
AC Q661U0;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Glycerol-3-phosphate O-acyltransferase, putative.
GN ORFNames=BG0328;
OS Borrelia garinii pfi.
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=290434;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Pfi;
RA Gloeckner G., Lehmann R., Romualdi A., Pradella S.,
RA Schulte-Spechtel U., Wilske B., Suehnel J., Platzer M.;
RT "Comparative analysis of the Borrelia garinii genome.";
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CP000013; AAU07181.1; -.
DR GO; GO:0008415; F:acyltransferase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR InterPro; IPR002123; Acyltransferase.
DR SMART; SM00563; Pf8c; 1.
KW Acyltransferase; Transferase.
SQ SEQUENCE 297 AA; 34455 MW; 77410BE6104D3A56 CRC64;

Query Match 82.8%; Score 24; DB 2; Length 297;
Best Local Similarity 57.1%; Pred. No. 5.4e+02;

```

```

Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
   : |||||
Db 244 VMSIAEF 250

RESULT 45
Q9U375 PRELIMINARY; PRT; 317 AA.
AC Q9U375;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein T08D2.2.
GN ORFNames=T08D2.2;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA McMurray A.A.;
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z92839; CAB07417.2; -.
DR WormBase; WBGene00011607; T08D2.2.
DR WormPep; T08D2.2; CE25108.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0003924; F:catalytic activity; IEA.
DR GO; GO:0006629; F:lipid metabolism; IEA.
DR InterPro; IPR001810; F-box.
DR InterPro; IPR000715; Glyco_trans_4.
DR Pfam; PF00646; F-box; 1.
DR Pfam; PF00953; Glycos_transf_4; 1.
KW Hypothetical protein.
SQ SEQUENCE 317 AA; 36591 MW; 145E88B02B75AE42 CRC64;

Query Match 82.8%; Score 24; DB 2; Length 317;
Best Local Similarity 83.3%; Pred. No. 5.7e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 MXVAEF 7
   : |||||
Db 223 MSVAEF 228

RESULT 46
K6PF_STRPN STANDARD; PRT; 335 AA.
AC Q97RC6;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE 6-phosphofructokinase (EC 2.7.1.11) (Phosphofructokinase)
GN Name=pfkA; OrderedLocusNames=SP0896;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC BAA-334 / TIGR4;
RX MEDLINE=21357209; PubMed=11463916; DOI=10.1126/science.1061217;

```

```

RA Tettelin H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,
RA Peterson S.N., Heidelberg J.F., DeBoy R.T., Haft D.H., Dodson R.J.,
RA Durkin A.S., Gwinn M.L., Kolonay J.F., Nelson W.C., Peterson J.D.,
RA Holtzapple E.K., Khouri H.M., Wolf A.M., Utterback T.R., Hansen C.L.,
RA McDonald L.A., Feldblyum T.V., Angluoli S.V., Dickinson T.,
RA Hickey E.K., Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.,
RA Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;
RT "Complete genome sequence of a virulent isolate of Streptococcus
RT pneumoniae";
RL Science 293:498-506(2001)
CC -1- CATALYTIC ACTIVITY: ATP + D-fructose 6-phosphate = ADP + D-
CC fructose 1,6-bisphosphate.
CC -1- PATHWAY: Key control step of glycolysis.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the phosphofructokinase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; AE007395; AAK75023.1; -.
DR PIR; F95103; F95103.
DR HSSP; P00512; 3PFK.
DR TIGR; SP0896; -.
DR HAMAP; MF_00339; -.
DR InterPro; IPR000023; Ppfruckinase.
DR Pfam; PF00365; PFK; 1.
DR PRINTS; PR00476; PHFRCTKINASE.
DR ProDom; PD000707; Ppfruckinase; 1.
DR PROSITE; PS00433; PHOSPHOFRUCTOKINASE; 1.
KW Allosteric enzyme; ATP-binding; Complete proteome; Glycolysis; Kinase;
KW Magnesium; Transferase.
FT NP_BIND 21 25 ATP (By similarity).
FT NP_BIND 154 158 ATP (By similarity).
FT NP_BIND 171 187 ATP (By similarity).
FT ACT_SITE 127 127 Proton acceptor (By similarity).
FT BINDING 162 162 Substrate (By similarity).
FT BINDING 244 244 Substrate (By similarity).
FT BINDING 250 250 Substrate (By similarity).
FT BINDING 253 253 Substrate (By similarity).
SQ SEQUENCE 335 AA; 35174 MW; 2DFBAAC55CAFE6 CRC64;

Query Match 82.8%; Score 24; DB 1; Length 335;
Best Local Similarity 71.4%; Pred. No. 6e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
Db 224 VMSAEF 230

RESULT 47
K6PF_STR6
ID K6PF_STR6 STANDARD; PRT; 335 AA.
AC Q8DQ85;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE 6-phosphofructokinase (EC 2.7.1.11) (Phosphofructokinase)
DE (Phosphohexokinase).
GN Names=pfkA; OrderedLocusNames=spr0796;
OS Streptococcus pneumoniae (strain ATCC BAA-255 / R6).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=171101;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21429245; PubMed=11544234;

```

```

RX DOI=10.1128/JB.183.19.5709-5717.2001;
RA Hoskins J., Alborn W.B. Jr., Arnold J., Blaszcak L.C., Burgett S.,
RA DeHoff B.S., Estrem S.T., Fritz L., Fu D.-J., Fuller W., Geringer C.,
RA Gilmour R., Glass J.S., Khoja H., Kraft A.R., Lagace R.E.,
RA LeBlanc D.J., Lee L.N., Lefkowitz E.J., Lu J., Matsushima P.,
RA McAnen S.M., McHenney M., McLeaster K., Mundy C.W., Niclas T.I.,
RA Norris F.H., O'Garra M., Peery R.B., Robertson G.T., Rockey P.,
RA Sun P.-M., Winkler M.B., Yang Y., Young-Bellido M., Zhao G.,
RA Zook C.A., Baltz R.H., Jaskunas S.R., Rosteck P.R. Jr., Skatrud P.L.,
RA Glass J.I.;
RT "Genome of the bacterium Streptococcus pneumoniae strain R6.";
RL J. Bacteriol. 183:5709-5717(2001).
CC -1- CATALYTIC ACTIVITY: ATP + D-fructose 6-phosphate = ADP + D-
CC fructose 1,6-bisphosphate.
CC -1- PATHWAY: Key control step of glycolysis.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the phosphofructokinase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; AE008456; AAK99600.1; -.
DR PIR; D97971; D97971.
DR HSSP; P00512; 3PFK.
DR HAMAP; MF_00339; -.
DR InterPro; IPR000023; Ppfruckinase.
DR Pfam; PF00365; PFK; 1.
DR PRINTS; PR00476; PHFRCTKINASE.
DR ProDom; PD000707; Ppfruckinase; 1.
DR PROSITE; PS00433; PHOSPHOFRUCTOKINASE; 1.
KW Allosteric enzyme; ATP-binding; Complete proteome; Glycolysis; Kinase;
KW Magnesium; Transferase.
FT NP_BIND 21 25 ATP (By similarity).
FT NP_BIND 154 158 ATP (By similarity).
FT NP_BIND 171 187 ATP (By similarity).
FT ACT_SITE 127 127 Proton acceptor (By similarity).
FT BINDING 162 162 Substrate (By similarity).
FT BINDING 244 244 Substrate (By similarity).
FT BINDING 250 250 Substrate (By similarity).
FT BINDING 253 253 Substrate (By similarity).
SQ SEQUENCE 335 AA; 35213 MW; 9303242307FFC6F CRC64;

Query Match 82.8%; Score 24; DB 1; Length 335;
Best Local Similarity 71.4%; Pred. No. 6e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
Db 224 VMSAEF 230

RESULT 48
Q88SF9
ID Q88SF9 PRELIMINARY; PRT; 336 AA.
AC Q88SF9;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Transcription regulator of beta-galactosidase gene.
DE Names=lacR; OrderedLocusNames=lp_3470;
OS Lactobacillus plantarum.
OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
OC Lactobacillus.
OX NCBI_TaxID=1590;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=NCIMB 8826 / WCFS1;
RX MEDLINE=22480296; PubMed=12566566; DOI=10.1073/pnas.0337704100;

```

```

RA Kleerebezem M., Boekhorst J., van Kranenburg R., Molenaar D.,
RA Kuipers O.P., Leer R., Turchini R., Peters S.A., Sandbrink H.M.,
RA Fiers M.W.E.J., Stiekema W., Klein Lankester R.M., Bron P.A.,
RA Hoffer S.M., Nierop Groot M.N., Kerkhoven R., De Vries M., Ursing B.,
RA De Vos W.M., Siezen R.J.;
RT "Complete genome sequence of Lactobacillus plantarum WCFS1.";
RL Proc. Natl Acad. Sci. U.S.A. 100:1990-1995(2003).
DR EMBL; AL935262; CAD65557.1; -.
DR HSSP; P15039; 1DBQ.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR000843; HTH_Lact.
DR InterPro; IPR010982; Lambda-like DNA.
DR Pfam; PF00356; LactI; 1.
DR Pfam; PF00532; Pexipla BP 1; 1.
DR PROSITE; PS0932; HTH_LACT_2; 1.
KW Complete proteome.
SQ SEQUENCE 336 AA; 37694 MW; 8DE5C36A11C2A419 CRC64;

Query Match 82.8%; Score 24; DB 2; Length 336;
Best Local Similarity 71.4%; Pred. No. 6.1e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 VMKVAEF 7
DB 111 VMAVGEF 117

RESULT 49
Q63TY5 PRELIMINARY; PRT; 337 AA.
AC Q63TY5;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Putative ribose transport system, permease protein.
GN ORFNAMES=BPSL1832;
OS Burkholderia pseudomallei K96243.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Burkholderia.
OX NCBI_TaxID=272560;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=K96243;
RX PubMed=15377794;
RA Holden M.T.G., Titball R.W., Peacock S.J., Cerdano-Tarraga A.M.,
RA Atkins T., Crossman L.C., Pitt T., Churcher C., Mungall K.,
RA Bentley S.D., Sebahia M., Thomson N.R., Bason N., Beacham I.R.,
RA Brooks K., Brown K.A., Brown N.F., Challis G.L., Cherevach I.,
RA Chillingworth T., Cronin A., Crosset B., Davis P., Deshazer D.,
RA Fellwell T., Fraser A., Hance Z., Hauser H., Holroyd S., Jagels K.,
RA Keith K.B., Maddison M., Moule S., Price C., Quail M.A.,
RA Rabinovitch E., Rutherford K., Sanders M., Simmonds M.,
RA Songvilai S., Stevens K., Tumpala S., Vesaratchavest M.,
RA Whitehead S., Yeats C., Barrall B.G., Oyston P.C.F., Parkhill J.;
RT "Genomic plasticity of the causative agent of melioidosis,
RT Burkholderia pseudomallei."
RL Proc. Natl. Acad. Sci. U.S.A. 101:14240-14245(2004).
DR EMBL; BX571965; CAH35831.1; -.
SQ SEQUENCE 337 AA; 34345 MW; 283FDD1DB18D947D CRC64;

Query Match 82.8%; Score 24; DB 2; Length 337;
Best Local Similarity 71.4%; Pred. No. 6.1e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMKVAEF 7
DB 193 VMVAQF 199

RESULT 50

```

```

Q97KM0 PRELIMINARY; PRT; 367 AA.
ID Q97KM0;
AC Q97KM0;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Fusion: chorismate mutase and shikimate 5-dehydrogenase.
GN OrderedLocusNames=CAC0897;
OS Clostridium acetobutylicum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1488;
[1]
RN [1] SEQUENCE FROM N.A.
RP STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
RX MEDLINE=21359325; PubMed=11466286;
RA DOI=10.1128/JB.183.16.4823-4838.2001;
RA Noelting J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q.,
RA Gibson R., Lee H.M., Dubois J., Qiu D., Hitti J., Wolf Y.I.,
RA Tatusov R.L., Sabathe F., Doucette-Stamm L.A., Soucaille P.,
RA Daly M.J., Bennett G.N., Koonin E.V., Smith D.R.;
RT "Genome sequence and comparative analysis of the solvent-producing
RT bacterium Clostridium acetobutylicum."
RL J. Bacteriol. 183:4823-4838(2001).
DR EMBL; AF007605; AAK78873.1; -.
DR PIR; F97010; F97010.
DR HSSP; Q58484; 1NVT.
DR GO; GO:0005737; C:cyctoplasm; IEA.
DR GO; GO:0004106; F:chorismate mutase activity; IEA.
DR GO; GO:0004764; F:shikimate 5-dehydrogenase activity; IEA.
DR GO; GO:0016089; P:aromatic amino acid family biosynthesis, sh. . ; IEA.
DR InterPro; IPR011342; AroE.
DR InterPro; IPR002701; Chorismate_mut.
DR InterPro; IPR011279; CM_mono_gmpms.
DR InterPro; IPR006151; Shikimate_DH.
DR Pfam; PF01817; CM_2; 1.
DR Pfam; PF01488; Shikimate_DH; 1.
DR TIGRPFAMS; TIGR00507; aroE; 1.
DR TIGRPFAMS; TIGR01805; CM_mono_gmpms; 1.
KW Complete proteome; Oxidoreductase.
SQ SEQUENCE 367 AA; 41989 MW; CEACFA0B77F14842 CRC64;

Query Match 82.8%; Score 24; DB 2; Length 367;
Best Local Similarity 71.4%; Pred. No. 6.6e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMKVAEF 7
DB 29 VMEVAEF 35

RESULT 51
Q7Q854 PRELIMINARY; PRT; 375 AA.
ID Q7Q854;
AC Q7Q854;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE AgCP15202 (Fragment).
GN Name=agCG50918; ORFNAMES=ENSANGG0000009886;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.
OX NCBI_TaxID=180454;
[1]
RN [1] SEQUENCE FROM N.A.
RP STRAIN=PEST.
RC Anopheles Genome Sequencing Consortium;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
CC -! CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAB01008944; EAA10268.1; -.

```


DR GO: 0009966; P: regulation of signal transduction; IEA.
DR GO: 0009607; P: response to biotic stimulus; IEA.
DR InterPro: IPR007304; TAP42.
DR Pfam: PF04177; TAP42; 1.
FT NON TER 1
SQ SEQUENCE 375 AA; 43422 MW; B320E106818EC77A CRC64;

Query Match 82.8%; Score 24; DB 2; Length 375;
Best Local Similarity 83.3%; Pred. No. 6.7e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 VMXAEF 7
Db 291 MTVAEF 296

RESULT 52

Q8L789 PRELIMINARY; PRT; 378 AA.
AC Q8L789;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein At5g06420.
GN Name=At5g06420;
OS Arabidopsis thaliana (Mouse-ear cross).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC euroids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Southwick A., Nguyen M., Tripp M., Palm C.J., Jones T., Wu T.,
RA Carninci P., Chen H., Cheuk R., Chan M.M., Chang C.H., Dale J.M.,
RA Deng J.M., Hayashizaki Y., Hsuan V.W., Lee J.M., Ishida J., Kamiya A.,
RA Kawai J., Kim C.J., Narusaka M., Quach H.L., Sakurai T., Satou M.,
RA Seki M., Shinn P., Tang C.C., Torumi M., Wallender E.K., Wong C.,
RA Wu H.C., Yamada K., Yu G., Yuan S., Shinozaki K., Ecker J.,
RA Theologis A., Davis R.W.;
RA Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
RL [2]
RP SEQUENCE FROM N.A.
RA Nguyen M., Karlin-Neumann G., Southwick A., Tripp M., Miranda M.,
RA Palm C.J., Bowler L., Jones T., Banh J., Carninci P., Chen H.,
RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,
RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,
RA Ecker J., Theologis A., Davis R.W.;
RA Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
CC -I- SIMILARITY: Contains 1 RING-type zinc finger.
DR EMBL; AY136406; ANM97072.1; -
DR EMBL; BT000229; AAN15548.1; -
DR GO: 0000151; C:ubiquitin ligase complex; IEA.
DR GO: 00003676; F:nucleic acid binding; IEA.
DR GO: 0004842; F:ubiquitin-protein ligase activity; IEA.
DR GO: 0008270; F:zinc ion binding; IEA.
DR GO: 0016567; P:protein ubiquitination; IEA.
DR InterPro; IPR000571; Znf_CCCH.
DR Pfam; PF00097; zf-C3HC4; 1.
DR SMART; SM00184; RING; 1.
DR PROSITE; PS00518; ZF_RING_1; UNKNOWN_1.
DR PROSITE; PS0089; ZF_RING_2; 1.
DR Hypothetical protein; Metal-binding; Zinc; Zinc-finger.
SQ SEQUENCE 378 AA; 42518 MW; 142E4A6534BECA4D CRC64;

Query Match 82.8%; Score 24; DB 2; Length 378;
Best Local Similarity 71.4%; Pred. No. 6.8e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VMXAEF 7

Db 48 VMXAEF 54

RESULT 53

Q9FNG6 PRELIMINARY; PRT; 378 AA.
AC Q9FNG6;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Similarity to zinc finger protein (Hypothetical protein).
OS Arabidopsis thaliana (Mouse-ear cross).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC euroids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9806011; PubMed=9405937;
RA Kotani H., Nakamura Y., Sato S., Kaneko T., Asamizu E., Miyajima N.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. II.
RT Sequence features of the regions of 1,044,062 bp covered by thirteen
RT physically assigned P1 clones.";
RL DNA Res. 4:291-300(1997).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=22088475; PubMed=12093376;
RA Haas B.J., Volfovsky N., Town C.D., Troukhan M., Alexandrov N.,
RA Feldmann K.A., Flavell R.B., White O., Salzberg S.L.;
RT "Full-length messenger RNA sequences greatly improve genome
RT annotation.";
RL Genome Biol. 3:RESEARCH0029-RESEARCH0029(2002).
RN [3]
RP SEQUENCE FROM N.A.
RA Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,
RA Feldmann K.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
CC -I- SIMILARITY: Contains 1 RING-type zinc finger.
DR EMBL; AB006700; BAB08964.1; -
DR EMBL; AY087435; AAM67329.1; -
DR GO: 0000151; C:ubiquitin ligase complex; IEA.
DR GO: 00003676; F:nucleic acid binding; IEA.
DR GO: 0004842; F:ubiquitin-protein ligase activity; IEA.
DR GO: 0008270; F:zinc ion binding; IEA.
DR GO: 0016567; P:protein ubiquitination; IEA.
DR InterPro; IPR000571; Znf_CCCH.
DR Pfam; PF00097; zf-C3HC4; 1.
DR SMART; SM00184; RING; 1.
DR PROSITE; PS00518; ZF_RING_1; UNKNOWN_1.
DR PROSITE; PS0089; ZF_RING_2; 1.
DR Hypothetical protein; Metal-binding; Zinc; Zinc-finger.
SQ SEQUENCE 378 AA; 42460 MW; 173D71BB88A3FE2D CRC64;

Query Match 82.8%; Score 24; DB 2; Length 378;
Best Local Similarity 71.4%; Pred. No. 6.8e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VMXAEF 7

Db 48 VMXAEF 54

RESULT 54

Q6K7P6 PRELIMINARY; PRT; 414 AA.
ID Q6K7P6
AC Q6K7P6;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)

```

DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein P0452F04.40.
GN Name=P0452F04.40;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA Sasaki T., Matsumoto T., Yamamoto K.;
RL SUBMITTED (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP004776; BAD23066.1; -.
DR InterPro; IPR008941; TPR-like.
KW Hypothetical protein.
SQ SEQUENCE 414 AA; 46194 MW; F7D292D41DF66FF8 CRC64;

Query Match      82.8%; Score 24; DB 2; Length 414;
Best Local Similarity 57.1%; Pred. No. 7.4e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
Db 200 VMTIADF 206

RESULT 55
Q67UR2
ID Q67UR2 PRELIMINARY; PRT; 419 AA.
AC Q67UR2;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE GCN5-related N-acetyltransferase-like.
GN Name=P0453H04.20;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA Sasaki T., Matsumoto T., Katayose Y.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 6, PAC
clone:P0453H04.";
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP005453; BAD38107.1; -.
DR GO; GO:0016740; F:transferase activity; IEA.
DR InterPro; IPR000182; GCN5acetyl trans.
DR Pfam; PF00583; Acetyltransf_1; 1.
KW Transferase.
SQ SEQUENCE 419 AA; 45695 MW; B125377DF01D82C3 CRC64;

Query Match      82.8%; Score 24; DB 2; Length 419;
Best Local Similarity 71.4%; Pred. No. 7.5e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
Db 94 VMLVAEY 100

RESULT 56
Q7M8F1
ID Q7M8F1 PRELIMINARY; PRT; 419 AA.
AC Q7M8F1;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE PREPROTEIN TRANSLOCASE SUBUNIT.
GN Name=SECY; OrderedLocusNames=WS1698;
OS Wolinella succinogenes.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;

```

```

OC Helicobacteraceae; Wolinella.
OX NCBI_TaxID=844;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DSMZ 1740;
RX MEDLINE=22882897; PubMed=1450908; DOI=10.1073/pnas.1932838100;
RA Baar C., Eppinger M., Raddatz G., Simon J., Lanz C., Klimmek O.,
RA Nandakumar R., Gross R., Rosinus A., Keller H., Jagtap P., Linke B.,
RA Meyer F., Lederer H., Schuster S.C.;
RT "Complete genome sequence and analysis of Wolinella succinogenes.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:11690-11695(2003).
CC -!- FUNCTION: Involved in protein export (By similarity).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -!- SIMILARITY: Belongs to the secY/SEC61-alpha family.
DR EMBL; BX571661; CAE10725.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0015450; F:protein translocase activity; IEA.
DR GO; GO:0009306; P:protein secretion; IEA.
DR GO; GO:0006605; P:protein targeting; IEA.
DR GO; GO:0015031; P:protein transport; IEA.
DR InterPro; IPR002208; SecY.
DR Pfam; PF00344; SecY; 1.
DR PRINTS; PR00303; SECYRNLCASE.
DR TIGRFAMs; TIGR00967; 3a0501s007; 1.
DR PROSITE; PS00755; SECY_1; 1.
DR PROSITE; PS00756; SECY_2; 1.
KW Complete proteome; Protein transport; Translocation; Transmembrane;
KW Transport.
KW SEQUENCE 419 AA; 45248 MW; 7CAAFCE2EA0FBCF5C CRC64;

Query Match      82.8%; Score 24; DB 2; Length 419;
Best Local Similarity 57.1%; Pred. No. 7.5e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
Db 285 VMAIADF 291

RESULT 57
Q8KI64
ID Q8KI64 PRELIMINARY; PRT; 422 AA.
AC Q8KI64;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Similar to UDP-glucose/GDP-mannose dehydrogenase.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22053227; PubMed=12057956;
RX DOI=10.1128/JB.184.13.3614-3622.2002;
RA Raymond C.K., Sims E.H., Kas A., Spencer D.H., Kutyavin T.V.,
RA Ivey R.G., Zhou Y., Kaul R., Clendenning J.B., Olson M.V.;
RT "Genetic variation at the O-antigen biosynthetic locus in Pseudomonas
aeruginosa.";
RL J. Bacteriol. 184:3614-3622(2002).
DR EMBL; AF498401; AAM27558.1; -.
DR EMBL; AF498411; AAM27709.1; -.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR008927; 6GDGH_C like.
DR Pfam; PF00984; UDPG_MGDP_dh; 1.
DR Pfam; PF03720; UDPG_MGDP_dh_C; 1.
DR Pfam; PF03721; UDPG_MGDP_dh_N; 1.
SQ SEQUENCE 422 AA; 46019 MW; 28885984FB1A88A7 CRC64;

Query Match      82.8%; Score 24; DB 2; Length 422;
Best Local Similarity 83.3%; Pred. No. 7.5e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```
QY      2 MXVAEF 7
      305 MVAEAF 310
Db
RESULT 58
Q7P4E5 PRELIMINARY; PRT; 425 AA.
AC Q7P4E5;
DT 01-MAR-2004 (TReMBLrel. 26, Created)
DT 01-MAR-2004 (TReMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE NAD-specific glutamate dehydrogenase (EC 1.4.1.2).
GN Names:FN0417;
OS Fusobacterium nucleatum subsp. vincentii ATCC 49256.
OC Bacteria; Fusobacteria; Fusobacteriales; Fusobacteriaceae;
OC Fusobacterium.
OX NCBI_TaxID=209882;
RN [1]
RC STRAIN=ATCC 49256;
RA Karpatal V., Ivanova N., Anderson I., Reznik G., Bhattacharyya A.,
RA Gardner W.L., Mikhailova N., Larsen N., D'Souza M., Walunas T.,
RA Haselkorn R., Overbeek R., Kyrpides N.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC EMBL; AABF01000120; EAA23508.1; -.
DR HSP; P80319; IOTM.
DR GO; GO:0004352; F:glutamate dehydrogenase activity; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006520; P:amino acid metabolism; IEA.
DR InterPro; IPR006095; GLFV_dehydrog.
DR InterPro; IPR006096; GLFV_dehydrog.C.
DR InterPro; IPR006097; GLFV_dehydrog_N.
DR Pfam; PF00208; GLFV_dehydrog_1.
DR Pfam; PF02812; GLFV_dehydrog_N; 1.
DR PRINTS; PR00082; GLFVDRGNASE.
DR PROSITE; PS00074; GLFV_DEHYDROGNASE; 1.
KW Oxidoreductase.
SQ SEQUENCE 425 AA; 46526 MW; 1269189AA6C03038 CRC64;

Query Match 82.8%; Score 24; DB 2; Length 425;
Best Local Similarity 71.4%; Pred. No. 7.6e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 VMXVAEF 7
      238 VVAEAF 244
Db
RESULT 59
Q6UD01 PRELIMINARY; PRT; 426 AA.
AC Q6UD01;
DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
DE Predicted metal-dependent amidase/aminoacylase/carboxypeptidase.
GN ORFNames=ANT32C12.25;
OS uncultured marine proteobacterium ANT32C12.
OC Bacteria; Proteobacteria; environmental samples.
OX NCBI_TaxID=248048;
RN [1]
RC SEQUENCE FROM N.A.
RA De La Torre J.R., Christianson L.M., Beja O., Suzuki M.T., Karl D.M.,
RA Heidelberg J., DeLong E.F.;
RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
DR GO; GO:0004180; F:carboxypeptidase activity; IEA.
DR GO; GO:0008237; F:metallopeptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR002933; Peptidase M20.
DR InterPro; IPR010168; Pept M20D_amidh.
DR Pfam; PF01546; Peptidase M20; 1.
DR TIGRFAMS; TIGR01891; amidohydrolases; 1.
KW Carboxypeptidase.
SQ SEQUENCE 426 AA; 46479 MW; FC983028A03E345B CRC64;

Query Match 82.8%; Score 24; DB 2; Length 426;
Best Local Similarity 71.4%; Pred. No. 7.6e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 VMXVAEF 7
      139 LMGVAEF 145
Db
RESULT 61
Q8RG30 PRELIMINARY; PRT; 439 AA.
ID Q8RG30
```

```
RN [2]
RP SEQUENCE FROM N.A.
RA de la Torre J.R., Christianson L.M., Beja O., Suzuki M.T., Karl D.M.,
RA Heidelberg J.F., DeLong E.F.;
RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY372453; AAR05239.1; -.
DR GO; GO:0004180; F:carboxypeptidase activity; IEA.
DR GO; GO:0008237; F:metallopeptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR002933; Peptidase M20.
DR InterPro; IPR010168; Pept M20D_amidh.
DR Pfam; PF01546; Peptidase M20; 1.
DR TIGRFAMS; TIGR01891; amidohydrolases; 1.
KW Carboxypeptidase.
SQ SEQUENCE 426 AA; 46506 MW; 28B6CFFA7F07F438 CRC64;

Query Match 82.8%; Score 24; DB 2; Length 426;
Best Local Similarity 71.4%; Pred. No. 7.6e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 VMXVAEF 7
      139 LMGVAEF 145
Db
RESULT 60
Q6UD29 PRELIMINARY; PRT; 426 AA.
ID Q6UD29;
AC Q6UD29;
DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
DE Predicted metal-dependent amidase/aminoacylase/carboxypeptidase.
GN ORFNames=ANT8C10.36;
OS uncultured marine proteobacterium ANT8C10.
OC Bacteria; Proteobacteria; environmental samples.
OX NCBI_TaxID=248047;
RN [1]
RC SEQUENCE FROM N.A.
RA De La Torre J.R., Christianson L.M., Beja O., Suzuki M.T., Karl D.M.,
RA Heidelberg J., DeLong E.F.;
RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY372452; AAR05211.1; -.
DR GO; GO:0004180; F:carboxypeptidase activity; IEA.
DR GO; GO:0008237; F:metallopeptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR002933; Peptidase M20.
DR InterPro; IPR010168; Pept M20D_amidh.
DR Pfam; PF01546; Peptidase M20; 1.
DR TIGRFAMS; TIGR01891; amidohydrolases; 1.
KW Carboxypeptidase.
SQ SEQUENCE 426 AA; 46479 MW; FC983028A03E345B CRC64;

Query Match 82.8%; Score 24; DB 2; Length 426;
Best Local Similarity 71.4%; Pred. No. 7.6e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 VMXVAEF 7
      139 LMGVAEF 145
Db
RESULT 61
Q8RG30 PRELIMINARY; PRT; 439 AA.
ID Q8RG30
```

AC QBRG30;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE NAD-specific glutamate dehydrogenase (BC 1.4.1.21).
 GN OrderedLocusNames=FN0488;
 OS Fusobacterium nucleatum (subsp. nucleatum).
 OC Bacteria; Fusobacteria; Fusobacteriales; Fusobacteriaceae;
 OC Fusobacterium.
 OX NCBI_TaxID=76856;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 25586;
 RX MEDLINE=21886394; PubMed=11889109;
 RX DOI=10.1128/JB.184.7.2005-2018.2002;
 RA Kapural V., Anderson I., Ivanova N., Reznik G., Los T., Lykidis A.,
 RA Bhattacharyya A., Bartman A., Gardner W., Grechkin G., Zhu L.,
 RA Vasieva O., Chu L., Kogan Y., Chaga O., Goltzman E., Bernal A.,
 RA Larsen N., D'Souza M., Walunas T., Pusch G., Haselkorn R.,
 RA Fongstein M., Kyrpides N.C., Overbeek R.;
 RT "Genome sequence and analysis of the oral bacterium Fusobacterium
 RT nucleatum strain ATCC 25586."
 RL J. Bacteriol. 184:2005-2018(2002).
 DR EMBL; AE010560; AAL94684.1; -;
 DR HSSP; P80319; 1CTM
 DR GO; GO:0004352; F:glutamate dehydrogenase activity; IEA.
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.
 DR GO; GO:0006520; P:amino acid metabolism; IEA.
 DR InterPro; IPR006095; GLFV dehydrog.
 DR InterPro; IPR006096; GLFV dehydrog.C.
 DR InterPro; IPR006097; GLFV dehydrog.N.
 DR Pfam; PF00208; GLFV dehydrog; 1.
 DR Pfam; PF02812; GLFV dehydrog.N; 1.
 DR PRINTS; PR00082; GLFDRHGNAAS.
 DR PROSITE; PS00074; GLFV_DEHYDROGENASE; 1.
 KW Complete proteome.
 SQ SEQUENCE 439 AA; 48243 MW; AC99923686713FD5 CRC64;
 Query Match 82.8%; Score 24; DB 2; Length 439;
 Best Local Similarity 71.4%; Pred. No. 7.8e+02;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 VMXVAEF 7
 Db 252 VMVAEF 258
 RESULT 62
 ID Q7XA19 PRELIMINARY; PRT; 442 AA.
 AC Q7XA19;
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Putative transposase-related protein.
 GN ORFNames=177013.38;
 OS Solanum bulbocastanum (wild potato).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
 OC lamids; Solanales; Solanaceae; Solanum.
 OX NCBI_TaxID=147425;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC PubMed=12872003; DOI=10.1073/pnas.1533501100;
 RX Song J., Bradeen J.M., Naess S.K., Raasch J.A., Wielgus S.M.,
 RA Haberlach G.T., Liu J., Kuang H., Austin-Phillips S., Buell C.R.,
 RA Helgeson J.P., Jiang J.;
 RT "Gene RB cloned from Solanum bulbocastanum confers broad spectrum
 RT resistance to potato late blight."
 RL Proc. Natl. Acad. Sci. U.S.A. 100:9128-9133(2003).
 DR EMBL; AY303171; AAP45187.1; -;
 DR InterPro; IPR004332; MUDR.
 DR InterPro; IPR000817; Prion.

DR Pfam; PF03108; MUDR; 1.
 SQ SEQUENCE 442 AA; 50695 MW; 214F0711137B56F1 CRC64;
 Query Match 82.8%; Score 24; DB 2; Length 442;
 Best Local Similarity 83.3%; Pred. No. 7.8e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 2 VMXVAEF 7
 Db 1 MTVAEF 6
 RESULT 63
 ID Q6ZS89 PRELIMINARY; PRT; 457 AA.
 AC Q6ZS89;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Hypothetical protein FLJ45732.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Oshima A., Takahashi-Fujii A., Tanase T., Imose N., Takeuchi K.,
 RA Arita M., Musashino K., Yuuki H., Hara H., Sugiyama T., Irie R.,
 RA Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,
 RA Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,
 RA Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K., Wagatsuma M.,
 RA Murakawa K., Kanehori K., Sugiyama A., Kawakami B., Suzuki Y.,
 RA Sugano S., Nagahari K., Masuho Y., Nagai K., Isogai T.;
 RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AK127634; BAC87065.1; -;
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0015662; F:ATPase activity, coupled to transmembrane m. . .; IEA.
 DR GO; GO:0006812; P:cation transport; IEA.
 DR InterPro; IPR001757; ATPase_E1-E2.
 DR PRINTS; PR00119; CATATPASE_
 SQ SEQUENCE 457 AA; 51438 MW; E197444DC2AF08A CRC64;
 Query Match 82.8%; Score 24; DB 2; Length 457;
 Best Local Similarity 71.4%; Pred. No. 8.1e+02;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 VMXVAEF 7
 Db 389 LMVVAEF 395
 RESULT 64
 ID Q8WZNS PRELIMINARY; PRT; 463 AA.
 AC Q8WZNS;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Putative ENA-ATPase (Fragment).
 GN Name=ena1;
 OS Pleurotus ostreatus (oyster mushroom) (White-rot fungus).
 OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;
 OC Agaricales; Pleurotaceae; Pleurotus.
 OX NCBI_TaxID=5322;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Benito B.;
 RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ420741; CAD12640.1; -;
 DR HSSP; P04191; 1SU4.
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0005524; F:ATP binding; IEA.

DR GO: 0015662; F:ATPase activity, coupled to transmembrane m. . . ; IEA.
 DR GO: 0003924; F: catalytic activity; IEA.
 DR GO: 0006812; P: cation transport; IEA.
 DR GO: 0008152; P: metabolism; IEA.
 DR Pfam: PF00702; Hydrolase; 1.
 DR PRINTS: PR00119; CATATPASE.
 DR TIGRFAMs: TIGR01494; ATPase_P-type; 1.
 FT NON_TER 1
 FT NON_TER 463
 SQ SEQUENCE 463 AA; 49188 MW; C38AE05528AAAFB28 CRC64;

Query Match 82.8%; Score 24; DB 2; Length 463;
 Best Local Similarity 71.4%; Pred. No. 8.2e+02;
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
 |||||
 DB 347 VMTAAEF 353

RESULT 65
 G64F DROME STANDARD; PRT; 469 AA.
 AC P83297; Q9VZJ5;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 05-JUL-2004 (Rel. 44, Last sequence update)
 DT 25-JAN-2005 (Rel. 46, Last annotation update)
 DE Putative gustatory receptor 64f.
 GN Name=Gr64f; ORFNames=CG32255;
 OS Drosophila melanogaster (fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkely;
 RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Anantides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abail J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
 RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreria S., Fleischmann W.,
 RA Folsler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D.A., Helman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lascko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Mout S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Moshrefi A.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Paclet J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun B.,
 RA Swirekas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,

RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).
 RN [2]
 RP GENOME REANNOTATION.
 RX MEDLINE=22426069; PubMed=12537572;
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
 RA Bettencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
 RA Scapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
 RA Lewis S.E.;
 RT "Annotation of the Drosophila melanogaster euchromatic genome: a
 RT systematic review";
 RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
 RN [3]
 RP IDENTIFICATION.
 RX MEDLINE=21407712; PubMed=11516643; DOI=10.1016/S0960-9822(01)00258-5;
 RA Dunipace L., Meister S., McNealy C., Amrein H.;
 RT "Spatially restricted expression of candidate taste receptors in the
 RT Drosophila gustatory system";
 RL Curr. Biol. 11:822-835(2001).
 CC -!- FUNCTION: Probable role in the gustatory response.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -!- SIMILARITY: Belongs to the G-protein coupled receptor Dr-tr
 CC family. Subfamily 11.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF003480; AAF47826.3; -
 CC FLYBase; FBgn0052255; Gr64f.
 DR GO: 0016021; C: integral to membrane; NAS.
 DR GO: 0008527; F: taste receptor activity; NAS.
 DR GO: 0050909; P: perception of taste; NAS.
 DR Pfam; PF06151; Trehalose recp; 1.
 KW G-protein coupled receptor; Hypothetical protein; Multigene family;
 KW Receptor; Transmembrane.
 FT DOMAIN 1 117 Extracellular (Potential).
 FT TRANSMEM 118 138 1 (Potential).
 FT DOMAIN 139 146 Cytoplasmic (Potential).
 FT TRANSMEM 147 167 2 (Potential).
 FT DOMAIN 168 199 Extracellular (Potential).
 FT TRANSMEM 200 220 3 (Potential).
 FT DOMAIN 221 265 Cytoplasmic (Potential).
 FT TRANSMEM 266 286 4 (Potential).
 FT DOMAIN 287 330 Extracellular (Potential).
 FT TRANSMEM 331 351 5 (Potential).
 FT DOMAIN 352 353 Cytoplasmic (Potential).
 FT TRANSMEM 354 374 6 (Potential).
 FT DOMAIN 375 435 Extracellular (Potential).
 FT TRANSMEM 436 456 7 (Potential).
 FT DOMAIN 457 469 Cytoplasmic (Potential).
 SQ SEQUENCE 469 AA; 54575 MW; E47B0232DF424AAA CRC64;

Query Match 82.8%; Score 24; DB 1; Length 469;
 Best Local Similarity 71.4%; Pred. No. 8.3e+02;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
 |||||
 DB 84 VLLVAEF 90

RESULT 66
 Q9C7R2
 ID Q9C7R2 PRELIMINARY; PRT; 484 AA.

AC 09C7R2;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Hypothetical protein F152.29.
 GN Names=F152.29;
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosid II; Brassicales; Brassicaceae; Arabidopses.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B.J., Wu D.,
 RA Waiti R., Renning C.M., Koo H., Fujii C.Y., Utterback T.R., Fraser C.M.,
 RA Barnstead M.E., Bowman C.L., White O., Nierman W.C., Fraser C.M.;
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Town C.D., Kaul S.;
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC068667; AG51739.1; -;
 DR PIR; E86416; E86416.
 DR InterPro; IPR008892; WCOB413.
 DR Pfam; PF05562; WCOB413; 2.
 DR Hypothetical protein
 KW Hypothetical protein
 SQ SEQUENCE 484 AA; 52066 MW; F94D3EA5C96059B3 CRC64;
 Query Match 82.8%; Score 24; DB 2; Length 484;
 Best Local Similarity 71.4%; Pred. No. 8.5e+02;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 VMXVAEF 7
 Db 437 VLOVAEF 443
 RESULT 67
 TP6B_AERPE STANDARD; PRT; 565 AA.
 AC Q9YEG4;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Type II DNA topoisomerase VI subunit B (EC 5.99.1.3) (TopoVI-B).
 GN Name=top6B; OrderedLocNames=APE0706;
 OS Aeropyrum pernix.
 OC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcales;
 OC Desulfurococaceae; Aeropyrum.
 OX NCBI_TaxID=56636;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=X1;
 RX MEDLINE=99310339; PubMed=10382966;
 RA Kawarabayasi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,
 RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Nishijima K., Nakazawa H.,
 RA Hosoyama A., Fukui S., Nagai Y., Otsuka R., Tanaka T., Kudoh Y.,
 RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,
 RA Yamazaki J., Kishida N., Oguchi A., Aoki K.-I., Kubota K.,
 RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
 RT "Complete genome sequence of an aerobic hyper-thermophilic
 crenarchaeon, Aeropyrum pernix X1.";
 RL DNA Res. 6:83-101(1999).
 CC -1- FUNCTION: Relaxes both positive and negative superturns and
 CC exhibits a strong decatenase activity. The B subunit binds ATP (By
 CC similarity).
 CC -1- CATALYTIC ACTIVITY: ATP-dependent breakage, passage and rejoining
 CC of double-stranded DNA.
 CC -1- SUBUNIT: Heterotetramer of two subunits A and two subunits B (By
 CC similarity).
 CC -1- SIMILARITY: Belongs to the TOP6B family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF000060; BAA79682.1; -;
 DR PIR; B72660; B72660.
 DR HSSP; O05207; IMX0.
 DR HAMAP; MF_00322; -; 1.
 DR InterPro; IPR003594; ATPbind_ATPase.
 DR InterPro; IPR005734; DNA_top6B.
 DR InterPro; IPR010979; Ribosomal_H2TH.
 DR Pfam; PF02518; HATPase_c; 1.
 DR SMART; SM00387; HATPase_c; 1.
 DR TIGRFAMs; TIGR01052; top6b; 1.
 DR ATP-binding; Complete proteome; DNA-binding; Isomerase; Topoisomerase.
 KW ATP-binding; Complete proteome; DNA-binding; Isomerase; Topoisomerase.
 SQ SEQUENCE 565 AA; 64096 MW; E4A3DFCC5E53D508 CRC64;
 Query Match 82.8%; Score 24; DB 1; Length 565;
 Best Local Similarity 83.3%; Pred. No. 9.9e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 2 MXVAEF 7
 Db 28 MSVAEF 33
 RESULT 68
 O58843 PRELIMINARY; PRT; 582 AA.
 ID O58843
 AC O58843;
 DT 01-AUG-1998 (TrEMBLrel. 07, Created)
 DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Hypothetical protein PH116.
 GN OrderedLocNames=PH116;
 OS Pyrococcus horikoshii.
 OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
 OC Pyrococcus.
 OX NCBI_TaxID=53953;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=OT3;
 RX MEDLINE=98344137; PubMed=9679194;
 RA Kawarabayasi Y., Sawada M., Horikawa H., Haikawa Y., Hino Y.,
 RA Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,
 RA Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohfuku Y.,
 RA Funahashi T., Tanaka T., Kudoh Y., Yamazaki J., Kishida N., Oguchi A.,
 RA Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,
 RA Masuchi Y., Shizuya H., Kikuchi H.;
 RT "Complete sequence and gene organization of the genome of a hyper-
 RT thermophilic archaeobacterium, Pyrococcus horikoshii OT3.";
 RL DNA Res. 5:55-76(1998).
 CC -1- FUNCTION: Exchanges the guanine residue with 7-aminomethyl-7-
 CC deazaguanine in tRNAs with GU(N) anticodons (tRNA-Asp, -Asn, -His
 CC and -Tyr). After this exchange, a cyclopentendiol moiety is
 CC attached to the 7-aminomethyl group of 7-deazaguanine, resulting
 CC in the hypermodified nucleoside queuosine (Q) (7-((4,5-cis-
 CC dihydroxy-2-cyclopenten-1-yl)amino)methyl)-7-deazaguanosine (By
 CC similarity).
 CC -1- CATALYTIC ACTIVITY: tRNA guanine + queuine = tRNA queuine +
 CC guanine.
 CC -1- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
 CC -1- SIMILARITY: Belongs to the queuine tRNA-ribosyltransferase family.
 DR EMBL; AF000005; BAA30215.1; -;
 DR PIR; E71052; E71052.
 DR PDB; 1IQ8; X-ray; A/B=1-582.
 DR PDB; 1IT7; X-ray; A/B=1-582.
 DR PDB; 1IT8; X-ray; A/B=1-582.
 DR PDB; 1J2B; X-ray; A/B=1-582.
 DR GO; GO:0008479; F:queuine tRNA-ribosyltransferase activity; IEA.

DR GO: 0003723; F:RNA binding; IEA.
 DR GO: 0016757; F:transferase activity, transferring glycosyl. . . ; IEA.
 DR GO: 0008616; P:queuosine biosynthesis; IEA.
 DR GO: 0008033; P:RNA processing; IEA.
 DR InterPro: IPR004804; tRNA_ribo_trans.
 DR InterPro: IPR002478; PUA.
 DR InterPro: IPR002616; tRNA_ribo_trans.
 DR InterPro: IPR004521; Unchar_dom_2.
 DR Pfam: PF01702; TGT; 1.
 DR SMART: SM00359; PUA; 1.
 DR TIGRFAMS: TIGR00432; arcsn_trna_tgt; 1.
 DR TIGRFAMS: TIGR00449; tgt_general; 1.
 DR TIGRFAMS: TIGR00451; unchar_dom_2; 1.
 DR PROSITE: PS00890; PUA; 1.
 DR Complete proteome; Glycosyltransferase; Hypothetical protein;
 KW Queuosine biosynthesis; Transferase; Zinc; tRNA processing.
 SQ SEQUENCE 582 AA; 66595 MW; B96FIDSEC0D73AC3 CRC64;

Query Match 82.8%; Score 24; DB 2; Length 582;
 Best Local Similarity 57.1%; Pred. No. 1e+03;
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEP 7
 ||| :|||
 Db 440 VMAIAEY 446

RESULT 69
 Q9UZN0 PRELIMINARY; PRT; 584 AA.

ID Q9UZN0
 AC Q9UZN0
 DT 01-MAY-2000 (T-EMBLrel. 13, Created)
 DT 01-MAY-2000 (T-EMBLrel. 13, Last sequence update)
 DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
 DE Queuine-archaeosine tRNA ribosyltransferase.
 GN ORFNames=PA80740;
 OS Pyrococcus abyssi.
 OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
 OC Pyrococcus.
 OX NCBI_TaxID=29292;
 RN [1]
 RC SEQUENCE FROM N.A.
 RP STRAIN=GE5 / Orsay;
 RX MEDLINE=22511545; PubMed=12622808;
 RA Cohen G.N., Barbe V., Flament D., Galperin M., Heilig R., Lecompte O.,
 RA Poch O., Prieur D., Querellou J., Ripp R., Thierry J.-C.,
 RA Van der Oost J., Weissenbach J., Zivanovic Y., Forterre P.;
 RT "An integrated analysis of the genome of the hyperthermophilic
 RT archaeon Pyrococcus abyssi.";
 RL Mol. Microbiol. 47:1495-1512(2003).
 CC -I- FUNCTION: Exchanges the guanine residue with 7-aminomethyl-7-
 CC deazaguanine in tRNAs with GU(N) anticodons (tRNA-Asp, -Asn, -His
 CC and -Tyr). After this exchange, a cyclopentendiol moiety is
 CC attached to the 7-aminomethyl group of 7-deazaguanine, resulting
 CC in the hypermodified nucleoside queuosine (Q) (7-((4,5-cis-
 CC dihydroxy-2-cyclopenten-1-yl)amino)methyl)-7-deazaguanosine) (By
 CC similarity).
 CC -I- CATALYTIC ACTIVITY: tRNA guanine + queuine = tRNA queuine +
 CC guanine.
 CC -I- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
 CC -I- SIMILARITY: Belongs to the queuine tRNA-ribosyltransferase family.
 DR EMBL: AJ248286; CAB50027.1; .
 DR FNR; F75090; F75090.
 DR HSP; OS8843; I1Q8.
 DR GO: 0008479; F:queuine tRNA-ribosyltransferase activity; IEA.
 DR GO: 0003723; F:RNA binding; IEA.
 DR GO: 0016757; F:transferase activity, transferring glycosyl. . . ; IEA.
 DR GO: 0008616; P:queuosine biosynthesis; IEA.
 DR GO: 0008033; P:RNA processing; IEA.
 DR InterPro: IPR004804; tRNA_ribo_trans.
 DR InterPro: IPR002478; PUA.
 DR InterPro: IPR002616; tRNA_ribo_trans.

DR InterPro: IPR004521; Unchar_dom_2.
 DR Pfam: PF01472; PUA; 1.
 DR Pfam: PF01702; TGT; 1.
 DR SMART: SM00359; PUA; 1.
 DR TIGRFAMS: TIGR00432; arcsn_trna_tgt; 1.
 DR TIGRFAMS: TIGR00449; tgt_general; 1.
 DR TIGRFAMS: TIGR00451; unchar_dom_2; 1.
 DR PROSITE: PS00890; PUA; 1.
 DR Complete proteome; Glycosyltransferase; Queuosine biosynthesis;
 KW Transferase; Zinc; tRNA processing.
 SQ SEQUENCE 584 AA; 66667 MW; E31755EAE3681666 CRC64;

Query Match 82.8%; Score 24; DB 2; Length 584;
 Best Local Similarity 57.1%; Pred. No. 1e+03;
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEP 7
 ||| :|||
 Db 440 VMAIAEY 446

RESULT 70
 Q8SRA9 PRELIMINARY; PRT; 619 AA.

ID Q8SRA9
 AC Q8SRA9
 DT 01-JUN-2002 (T-EMBLrel. 21, Created)
 DT 01-JUN-2002 (T-EMBLrel. 21, Last sequence update)
 DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
 DE ATP DEPENDENT DNA BINDING HELICASE (RAD3/XPD SUBFAMILY OF
 DE HELICASES).
 GN Name=ECU08_1120;
 OS Encephalitozoon cuniculi GB-M1.
 OC Eukaryota; Fungi; Microsporidia; Unikaryonidae; Encephalitozoon.
 OX NCBI_TaxID=284813;
 RN [1]
 RC SEQUENCE FROM N.A.
 RP STRAIN=GB-M1;
 RX MEDLINE=21576510; PubMed=11719806; DOI=10.1038/35106579;
 RA Katinka M.D., Duprat S., Cornillot E., Metenier G., Thomarar P.,
 RA Prensier G., Barbe V., Peyretailade E., Brottier P., Wincker P.,
 RA Delbac F., El Alaoui H., Peyret P., Saurin W., Gouy M.,
 RA Weissenbach J., Vivares C.P.;
 RT "Genome sequence and gene compaction of the eukaryote parasite
 RT Encephalitozoon cuniculi.";
 RL Nature 414:450-453(2001).
 RN [2]
 RC SEQUENCE FROM N.A.
 RP STRAIN=GB-M1;
 RA Genoscope;
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AL590448; CAD26418.1; .
 DR GO: 0005634; C:nucleus; IEA.
 DR GO: 0005524; F:ATP binding; IEA.
 DR GO: 0004003; F:ATP-dependent DNA helicase activity; IEA.
 DR GO: 0003677; F:DNA binding; IEA.
 DR GO: 0016818; F:hydrolase activity, acting on acid anhydrid. . . ; IEA.
 DR GO: 0006139; P:nucleobase, nucleoside, nucleotide and nucl. . . ; IEA.
 DR GO: 0005289; P:nucleotide-excision repair; IEA.
 DR InterPro: IPR010614; DEAD 2.
 DR InterPro: IPR002464; DEAH box.
 DR InterPro: IPR006554; DEXDC2.
 DR InterPro: IPR006555; Helic c2.
 DR InterPro: IPR001865; Ribosomal S2.
 DR InterPro: IPR001945; XPD_DNA_repair.
 DR Pfam: PF06733; DEAD 2; 1.
 DR SMART: SM00488; DEXDC2; 1.
 DR SMART: SM00491; HELICc2; 1.
 DR TIGRFAMS: TIGR00604; rad3; 1.
 DR PROSITE: PS00690; DEAH ATP HELICASE; 1.
 DR PROSITE: PS00962; RIBOSOMAL_S2_1; UNKNOWN_1.
 KW Helicase.
 SQ SEQUENCE 619 AA; 69072 MW; 9574F09FD2B52A67 CRC64;

Query Match 82.8%; Score 24; DB 2; Length 619;
 Best Local Similarity 71.4%; Pred. No. 1.1e+03;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
 |||
 Db 284 VMGVSEF 290

RESULT 71

Q8P8F1 PRELIMINARY; PRT; 689 AA.
 ID Q8P8F1
 AC Q8P8F1
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Hypothetical protein XCC2291.
 GN OrderedLocusNames=XCC2291;
 OS Xanthomonas campestris (pv. campestris).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
 OC Xanthomonadaceae; Xanthomonas.
 OX NCBI_TaxID=340;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 33913 / NCPPB 528;
 RX MEDLINE=22022145; PubMed=12024217; DOI=10.1038/417459a;
 RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
 RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
 RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
 RA Camarotte G., Cannavan F., Cardoso J., Chambergo F., Ciapina L.P.,
 RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorry H.,
 RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
 RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
 RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
 RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
 RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
 RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
 RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
 RA Spinola L.A.F., Takita M.A., Tamura R.B., Teixeira E.C., Tezza R.I.D.,
 RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
 RA Setubal J.C., Kitajima J.P.;
 RT "Comparison of the genomes of two Xanthomonas pathogens with differing
 RT host specificities.";
 RL Nature 417:459-463(2002).
 DR EMBL; AE012336; AW41570.1; -;
 DR GO; GO:0004871; F:signal transducer activity; IEA.
 DR GO; GO:0007165; P:signal transduction; IEA.
 DR InterPro; IPR002086; Aldehyd dehydrog.
 DR InterPro; IPR011006; Chey_like.
 DR InterPro; IPR001633; EAL.
 DR InterPro; IPR000160; GGDEF.
 DR InterPro; IPR000014; PAS.
 DR Pfam; PF00563; EAL; 1.
 DR Pfam; PF00990; GGDEF; 1.
 DR SMART; SM00052; DUF2; 1.
 DR SMART; SM00091; PAS; 1.
 DR TIGRFAMS; TIGR00254; GGDEF; 1.
 DR TIGRFAMS; TIGR00229; sensory box; 1.
 DR PROSITE; PS00687; ALDHDEHYD_DEHYDR_GLU; UNKNOWN_1.
 DR PROSITE; PS00883; EAL; 1.
 DR PROSITE; PS00887; GGDEF; 1.
 KW Complete proteome.
 SQ SEQUENCE 689 AA; 76025 MW; C62FD2B62C18E793 CRC64;

Query Match 82.8%; Score 24; DB 2; Length 689;
 Best Local Similarity 57.1%; Pred. No. 1.2e+03;
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
 ::|||
 Db 648 ILTVAEF 654

RESULT 72

P91063 PRELIMINARY; PRT; 738 AA.
 ID P91063
 AC P91063
 DT 01-MAY-1997 (TrEMBLrel. 03, Created)
 DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Hypothetical protein C17H11.2;
 GN Names=C17H11.2; ORFNames=C17H11.2;
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RX MEDLINE=99069613; PubMed=9851916;
 RG WormBase Consortium;
 RT "Genome sequence of the nematode C. elegans: a platform for
 RT investigating biology. The C. elegans Sequencing Consortium.";
 RL Science 282:2012-2018(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Johnson D.;
 RT "The sequence of C. elegans cosmid C17H11.";
 RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Waterston R.;
 RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RG WormBase Consortium;
 RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U80847; AAB37983.2; -;
 DR PIR; T25551; T25551.
 DR WormBase; WBGene00015922; C17H11.2.
 DR WormPep; C17H11.2; CB27703.
 KW Hypothetical protein.
 SQ SEQUENCE 738 AA; 83764 MW; D41414064A4944F2 CRC64;

Query Match 82.8%; Score 24; DB 2; Length 738;
 Best Local Similarity 83.3%; Pred. No. 1.3e+03;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 MXVAEF 7
 |||||
 Db 540 MTVAEF 545

RESULT 73

O846V6 PRELIMINARY; PRT; 871 AA.
 ID O846V6
 AC O846V6;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Isoleucyl-tRNA synthetase.
 GN Name=ileS;
 OS Staphylococcus epidermidis.
 OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
 OX NCBI_TaxID=1282;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22502866; PubMed=12615863;
 RA Yun H.J., Lee S.W., Yoon G.M., Kim S.Y., Choi S., Lee Y.S., Choi E.C.,
 RA Kim S.;
 RT "Prevalence and mechanisms of low- and high-level mupirocin resistance
 RT in staphylococci isolated from a Korean hospital.";
 RL J. Antimicrob. Chemother. 51:619-623(2003).


```

DR EMBL; AF516209; AAO65848.1; -.
DR HSSP; P41972; 1FFY.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004822; F:isoleucine--trNA ligase activity; IEA.
DR GO; GO:0006428; P:isoleucyl--trNA aminoacylation; IEA.
DR InterPro; IPR002300; trNA-synt_1a.
DR InterPro; IPR001412; trNA-synt_1.
DR InterPro; IPR002301; trNA-synt_1le.
DR InterPro; IPR009080; trNAeyn_1a_bind.
DR InterPro; IPR009008; ValRS_1leRS_edit.
DR Pfam; PF00133; trNA-synt_1; 1.
DR PRINTS; PR00984; TRNASYNTHILE.
DR TIGRFAMs; TIGR00392; iles; 1.
DR PROSITE; PS00178; AA_TRNA_LIGASE_1; 1.
KW Aminoacyl-tRNA synthetase; ATP-binding; Complete proteome; Ligase;
KW Metal-binding; Protein biosynthesis; Zinc.
FT SITE 57 67 "HIGH" region.
FT SITE 595 599 ATP (By similarity).
FT BINDING 598 598 ATP (By similarity).
SQ SEQUENCE 916 AA; 105185 MW; C5155509A91FCEE CRC64;

Query Match 82.8%; Score 24; DB 1; Length 916;
Best Local Similarity 83.3%; Pred. No. 1.6e+03;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 MVXAEF 7
Db 115 MSVAEF 120

RESULT 74
SYI_STAEP STANDARD; PRT; 916 AA.
ID SYI_STAEP STANDARD; PRT; 916 AA.
AC QCSXI;
DT 29-MAR-2004 (Rel. 43, Created)
DT 29-MAR-2004 (Rel. 43, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Isoleucyl-tRNA synthetase (EC 6.1.1.5) (Isoleucine--trNA ligase)
DE (IleRS).
GN Name=iles; OrderedLocusNames=SE0868;
OS Staphylococcus epidermidis.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1282;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 12228;
RX PubMed=12950922;
RA Zhang Y.-Q., Ren S.-X., Li H.-L., Wang Y.-X., Fu G., Yang J.,
RA Qin Z.-Q., Miao Y.-G., Wang W.-Y., Chen R.-S., Shen Y., Chen Z.,
RA Yuan Z.-H., Zhao G.-P., Qu D., Danchin A., Wen Y.-M.;
RT "Genome-based analysis of virulence genes in a non-biofilm-forming
RT Staphylococcus epidermidis strain (ATCC 12228).";
RL Mol. Microbiol. 49:1577-1593(2003).
CC -!- CATALYTIC ACTIVITY: ATP + L-isoleucine + trNA(Ile) = AMP +
CC diphosphate + L-isoleucyl-tRNA(Ile).
CC -!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
CC -!- SUBUNIT: Monomer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the class-I aminoacyl-tRNA synthetase
CC family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE016746; AAO04465.1; -.
DR HSSP; P41972; 1FFY.
DR InterPro; IPR002300; trNA-synt_1a.
DR InterPro; IPR001412; trNA-synt_1.
DR InterPro; IPR002301; trNA-synt_1le.
DR InterPro; IPR009080; trNAeyn_1a_bind.
DR InterPro; IPR009008; ValRS_1leRS_edit.
DR InterPro; IPR010663; 2F-PFG_IleRS.

```

```

DR Pfam; PF00133; trNA-synt_1; 1.
DR Pfam; PF06827; zf-PFG_IleRS; 1.
DR PRINTS; PR00984; TRNASYNTHILE.
DR TIGRFAMs; TIGR00392; iles; 1.
DR PROSITE; PS00178; AA_TRNA_LIGASE_1; 1.
KW Aminoacyl-tRNA synthetase; ATP-binding; Complete proteome; Ligase;
KW Metal-binding; Protein biosynthesis; Zinc.
FT SITE 57 67 "HIGH" region.
FT SITE 595 599 ATP (By similarity).
FT BINDING 598 598 ATP (By similarity).
SQ SEQUENCE 916 AA; 105185 MW; C5155509A91FCEE CRC64;

Query Match 82.8%; Score 24; DB 1; Length 916;
Best Local Similarity 83.3%; Pred. No. 1.6e+03;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 MVXAEF 7
Db 115 MSVAEF 120

RESULT 75
SYI_BACSU STANDARD; PRT; 921 AA.
ID SYI_BACSU STANDARD; PRT; 921 AA.
AC Q45477; O31730; P71022;
DT 01-NOV-1997 (Rel. 35, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Isoleucyl-tRNA synthetase (EC 6.1.1.5) (Isoleucine--trNA ligase)
DE (IleRS).
GN Name=iles; OrderedLocusNames=BSU15430;
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377; DOI=10.1038/36786;
RA Kunst F., Ogasawara N., Moser I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA Borries R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,
RA Entiaz K.-D., Errington J., Fabret C., Ferrati E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Ghim S.Y., Glaser P., Goffeau A., Gollightly E.J., Grandi G.,
RA Guiseppe G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M.,
RA Jones L.-M., Joris B., Karamata D., Kasahara Y., Klaert-Blanchard M.,
RA Klein C., Kobayashi Y., Koetter P., Koningsstein G., Krogh S.,
RA Kuman M., Kurita K., Lapidus A., Lardinois S., Lauber J.,
RA Lazarevic V., Lee S.M., Levine A., Liu H., Masuda S., Mauel C.,
RA Medigue C., Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S.,
RA Noback M., Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B.,
RA Park S.H., Parro V., Pohl T.M., Portetelle D., Porwollik S.,
RA Prescott A.M., Presecan E., Pujic P., Purnelle B., Rapoport G.,
RA Rey M., Reynolds S., Rieger M., Rivolta C., Rocha E., Roche R.,
RA Rose M., Sadaie Y., Sato T., Scanlan E., Schleich S., Schroeter R.,
RA Scoffone F., Sekiguchi J., Sekowska A., Seror S.J., Seror P.,
RA Shin B.S., Soldo B., Sorokin A., Tacconi E., Takagi T., Takahashi H.,
RA Takenaru K., Takeuchi M., Tamakoshi A., Tanaka T., Terpetra P.,
RA Tognoni A., Toato V., Uchiyama S., Vandenbol M., Vannier P.,
RA Vassarotti A., Viari A., Wambutt R., Wedler E., Wedler H.,
RA Weitzengger T., Winkers P., Wipat A., Yamamoto H., Yamane K.,
RA Yasumoto K., Yata K., Yoshida K., Yoshikawa H.F., Zumbstein E.,
RA Yoshikawa H., Danchin A.;
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
RT subtilis.";
RL Nature 390:249-256(1997).
RN [2]
RP SEQUENCE OF 1-121 FROM N.A.
RC STRAIN=168;

```

```
RA Stewart G.C., Cha J.H.;
RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 733-921 FROM N.A.
RC STRAIN=168;
RA Pragai Z., Tjalsma H., Bolhuis A., van Dijk J.M., Venema G., Bron S.;
RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
RC -1- CATALYTIC ACTIVITY: ATP + L-isoleucine + tRNA(Ile) = AMP +
CC diphosphate + L-isoleucyl-tRNA(Ile).
CC -1- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
CC -1- SUBUNIT: Monomer (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: Belongs to the class-I aminoacyl-tRNA synthetase
CC family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z99112; CAB13417.1; -
CC EMBL; U60901; AAB49280.1; -
CC EMBL; U48870; AAB57764.1; -
CC PIR; H69643; H69643.
CC HSSP; P41972; 1OU3.
CC Subtilisin; BG11792; iles.
CC InterPro; IPR002300; tRNA-synt 1a.
CC InterPro; IPR001412; tRNA-synt_1.
CC InterPro; IPR002301; tRNA-synt_1e.
CC InterPro; IPR009080; tRNA-syn_1a_bind.
CC InterPro; IPR009008; ValRS_IleRS_edit.
CC InterPro; IPR010663; ZF-PFG_IleRS.
CC Pfam; PF00133; tRNA-synt 1; 1.
CC Pfam; PF06827; zf-PFG_IleRS; 1.
CC PRINTS; PR00984; TRNASYNTHILE.
CC TIGRfams; TIGR00392; iles; 1.
CC PROSITE; PS00178; AA_TRNA_LIGASE_I; 1.
CC Aminoacyl-tRNA synthetase; ATP-binding; Complete proteome; Ligase;
CC Metal-binding; Protein biosynthesis; Zinc.
CC SITE 57 67 "HIGH" region.
CC SITE 594 598 "KMSKS" region.
CC BINDING 597 597 ATP (By similarity).
CC CONFLICT 744 744 S -> F (in Ref. 3).
CC SEQUENCE 921 AA; 104784 MW; 011F31E5F7460D43 CRC64;

Query Match 82.8%; Score 24; DB 1; Length 921;
Best Local Similarity 83.3%; Pred. No. 1.6e+03;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 MXVAEF 7
DB 116 MSVAEF 121

RESULT 76
Q636D1 PRELIMINARY; PRT; 921 AA.
AC Q636D1
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DE Isoleucine--tRNA ligase (isoleucyl-tRNA synthetase) (EC 6.1.1.5).
GN Names=iles; ORFNames=B7ZK3654;
OS Bacillus cereus ZK.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=286861;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ZK;
RA Brettin T.S., Bruce D., Challacombe J.F., Gilna P., Han C., Hill K.,
```

```
RA Hitchcock P., Jackson P., Keim P., Longmire J., Lucas S., Okinaka R.,
RL Richardson P., Rubin E., Tice H.;
RT "Complete genome sequence of Bacillus cereus ZK.";
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CP000001; AAU16613.1; -.
KW Aminoacyl-tRNA synthetase; Ligase.
SQ SEQUENCE 921 AA; 104620 MW; FBF9656AB45D53BB CRC64;

Query Match 82.8%; Score 24; DB 2; Length 921;
Best Local Similarity 83.3%; Pred. No. 1.6e+03;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 MXVAEF 7
DB 115 MTVAEF 120

RESULT 77
Q732H4 PRELIMINARY; PRT; 921 AA.
ID Q732H4
AC Q732H4
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Isoleucyl-tRNA synthetase (EC 6.1.1.5).
GN Names=iles; OrderedLocustNames=BCE3940;
OS Bacillus cereus (strain ATCC 10987).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=222523;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=14960714; DOI=10.1093/nar/gkh258;
RA Rasko D.A., Ravel J., Oekstad O.A., Helgason E., Cer R.Z., Jiang L.,
RA Shores K.A., Fouts D.E., Tourasse N.J., Angiuoli S.V., Kolonay J.F.,
RA Nelson W.C., Kolstoe A.-B., Fraser C.M., Read T.D.;
RT "The genome sequence of Bacillus cereus ATCC 10987 reveals metabolic
RT adaptations and a large plasmid related to Bacillus anthracis px01.";
RL Nucleic Acids Res. 32:977-988(2004).
DR EMBL; AE017276; AAS42843.1; -.
DR HSSP; P56690; 1JZS.
DR TIGR; BCE3940; -.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004822; F:isoleucine-tRNA ligase activity; IEA.
DR GO; GO:0006428; P:isoleucyl-tRNA aminoacylation; IEA.
DR InterPro; IPR002300; tRNA-synt_1a.
DR InterPro; IPR001412; tRNA-synt_1.
DR InterPro; IPR002301; tRNA-synt_1e.
DR InterPro; IPR009080; tRNA-syn_1a_bind.
DR InterPro; IPR010663; ZF-PFG_IleRS_edit.
DR Pfam; PF00133; tRNA-synt_1; 1.
DR Pfam; PF06827; zf-PFG_IleRS; 1.
DR PRINTS; PR00984; TRNASYNTHILE.
DR TIGRfams; TIGR00392; iles; 1.
DR PROSITE; PS00178; AA_TRNA_LIGASE_I; 1.
CC Complete proteome.
KW SEQUENCE 921 AA; 104589 MW; B531AF210A8036C9 CRC64;

Query Match 82.8%; Score 24; DB 2; Length 921;
Best Local Similarity 83.3%; Pred. No. 1.6e+03;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 MXVAEF 7
DB 115 MTVAEF 120

RESULT 78
Q819R4 PRELIMINARY; PRT; 921 AA.
ID Q819R4
AC Q819R4;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
```

DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Isoleucyl-tRNA synthetase (EC 6.1.1.5).
GN OrderedLocusNames=BC3895;
OS Bacillus cereus (strain ATCC 14579 / DSM 31).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=226900;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22608415; PubMed=12721630; DOI=10.1038/nature01582;
RA Ivanova N., Sorokin A., Anderson I., Galleron N., Candelon B.,
Kapural V., Bhattacharya A., Reznik G., Mikhailova N., Lapidus A.,
Chu L., Mazur M., Goltzman E., Larsen N., D'Souza M., Walunas T.,
Grecklin Y., Pusch G., Haselkorn R., Fonstein M., Ehrlich S.D.,
Overbeek R., Kyrpides N.C.;
RT "Genome sequence of Bacillus cereus and comparative analysis with
RT Bacillus anthracis.";
RL Nature 423:81-86(2003).
DR HSSP; P41972; 1FFY.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004822; F:isoleucine-tRNA ligase activity; IEA.
DR GO; GO:0016874; F:ligase activity; IEA.
DR GO; GO:0006428; P:isoleucyl-tRNA aminoacylation; IEA.
DR InterPro; IPR002300; tRNA-synt_1a.
DR InterPro; IPR001412; tRNA-synt_1.
DR InterPro; IPR002301; tRNA-synt_ile.
DR InterPro; IPR009080; tRNA-synt_1a_bind.
DR InterPro; IPR009008; ValRS ITERS edit.
DR InterPro; IPR010663; Zf-PFG_1IERS.
DR Pfam; PF00133; tRNA-synt_1; 1.
DR Pfam; PF06827; zf-PFG_1IERS; 1.
DR PRINTS; PR00984; TRNASYNTHILE.
DR TIGRFAMS; TIGR00392; iles; 1.
DR PROSITE; PS00178; AA-TRNA_LIGASE_I; 1.
KW Aminoacyl-tRNA synthetase; Complete proteome; Ligase.
SQ SEQUENCE 921 AA; 104561 MW; 522D045A316EE1E0 CRC64;

Query Match 82.8%; Score 24; DB 2; Length 921;
Best Local Similarity 83.3%; Pred. No. 1.6e+03;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 MVVAEF 7
Db 115 MTVAEF 120

RESULT 79
ID Q81WE4 PRELIMINARY; PRT; 921 AA.
AC Q81WE4; Q6HUJ1; Q6KNS7;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Isoleucyl-tRNA synthetase.
GN Names=iles-2; OrderedLocusNames=BA4034, BAS3746, GBAA4034;
OS Bacillus anthracis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1392;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Ames / isolate Porton;
RC MEDLINE=22608414; PubMed=12721629; DOI=10.1038/nature01586;
RA Read T.D., Peterson S.N., Tourasse N.J., Baillie L.W., Paulsen I.T.,
Nelson K.E., Tettelin H., Fouts D.E., Eisen J.A., Gill S.R.,
Holtzapple E.K., Ostad O.A., Helgason E., Ristone J., Wu M.,
Kolony J.F., Beaman M.J., Dodson R.J., Brinkac L.M., Gwinn M.L.,
DeBoy R.T., Madpu R., Daugherty S.C., Durkin A.S., Haft D.H.,
Nelson W.C., Peterson J.D., Pop M., Khouri H.M., Radune D.,
Benton J.L., Mahamoud Y., Jiang L., Hance I.R., Weidman J.F.,
Berry K.J., Plaut R.D., Wolf A.M., Watkins K.E., Nierman W.C.,
Hazen A., Cline R.T., Redmond C., Thwaite J.E., White O.,
Salzberg S.L., Thomson B., Friedlander A.M., Koehler T.M.,
Hanna P.C., Kolstoe A.-B., Fraser C.M.;

RT "The genome sequence of Bacillus anthracis Ames and comparison to
RT closely related bacteria.";
RL Nature 423:81-86(2003).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Ames / isolate 0581;
RA Ravel J., Rasko D.A., Shumway M.F., Jiang L., Cer R.Z., Federova N.B.,
Wilson M., Stanley S., Decker S., Read T.D., Salzberg S.L.,
Fraser C.M.;
RT "Bacillus anthracis comparative genomics.";
RL Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Sterne;
RA Brettin T.S., Bruce D., Challacombe J.F., Gilna P., Han C., Hill K.,
Hitchcock P., Jackson P., Keim P., Longmire J., Lucas S., Okinaka R.,
Richardson P., Rubin E., Tice H.;
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE017036; AAP27760.1; -.
DR EMBL; AE017334; AAT33151.1; -.
DR EMBL; AE017225; AAT56048.1; -.
DR HSSP; P41972; 1FFY.
DR TIGR; BA4034; -.
DR TIGR; GBAA4034; -.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004822; F:isoleucine-tRNA ligase activity; IEA.
DR GO; GO:0006428; P:isoleucyl-tRNA aminoacylation; IEA.
DR InterPro; IPR002300; tRNA-synt_1a.
DR InterPro; IPR001412; tRNA-synt_1.
DR InterPro; IPR002301; tRNA-synt_ile.
DR InterPro; IPR009080; tRNA-synt_1a_bind.
DR InterPro; IPR009008; ValRS ITERS edit.
DR InterPro; IPR010663; Zf-PFG_1IERS.
DR Pfam; PF00133; tRNA-synt_1; 1.
DR Pfam; PF06827; zf-PFG_1IERS; 1.
DR PRINTS; PR00984; TRNASYNTHILE.
DR TIGRFAMS; TIGR00392; iles; 1.
DR PROSITE; PS00178; AA-TRNA_LIGASE_I; 1.
KW Aminoacyl-tRNA synthetase; Complete proteome.
SQ SEQUENCE 921 AA; 104610 MW; 4B383BCCC606348B CRC64;

Query Match 82.8%; Score 24; DB 2; Length 921;
Best Local Similarity 83.3%; Pred. No. 1.6e+03;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 MVVAEF 7
Db 115 MTVAEF 120

RESULT 80
Q9K9V0 PRELIMINARY; PRT; 921 AA.
AC Q9K9V0;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Isoleucyl-tRNA synthetase.
GN Names=iles; OrderedLocusNames=BH2545;
OS Bacillus halodurans.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=86665;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=C-125;
RC MEDLINE=20512582; PubMed=11058132; DOI=10.1093/nar/28.21.4317;
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
Fuji F., Hikama C., Nakamura Y., Ogasawara N., Kuhara S.,
Horikoshi K.;
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT halodurans and genomic sequence comparison with Bacillus subtilis.";
RL Nucleic Acids Res. 28:4317-4331(2000).
DR EMBL; AP001515; BAB06264.1; -.

DR PIR; A83968; A83968.
 DR HSSP; P41972; 1FFV.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0004822; P:isoleucine-tRNA ligase activity; IEA.
 DR GO; GO:0006428; P:isoleucyl-tRNA aminoacylation; IEA.
 DR InterPro; IPR002300; tRNA-synt la.
 DR InterPro; IPR001412; tRNA-synt-1.
 DR InterPro; IPR002301; tRNA-synt-ile.
 DR InterPro; IPR009080; tRNA-synt-ile.
 DR InterPro; IPR009080; tRNA-synt-ile.
 DR InterPro; IPR010663; zf-PFG ilers.
 DR Pfam; PF001133; tRNA-synt 1; 1.
 DR Pfam; PF06827; zf-PFG ilers; 1.
 DR PRINTS; PR00984; tRNA-synt-1.
 DR TIGRFAMs; TIGR00392; iles; 1.
 DR PROSITE; PS00178; AA_TRNA_LIGASE I; 1.
 KW Aminoacyl-tRNA synthetase; Complete proteome.
 SQ SEQUENCE 921 AA; 104562 MW; 5D9A17BD5A741B9A CRC64;

Query Match 82.8%; Score 24; DB 2; Length 921;
 Best Local Similarity 83.3%; Pred. No. 1.6e+03;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 MKVAEP 7
 | | | | |
 Db 115 MSVAEP 120

RESULT 81

Q6HER9 PRELIMINARY; PRT; 921 AA.
 ID Q6HER9
 AC Q6HER9
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Isoleucine-tRNA ligase [Isoleucyl-tRNA synthetase] (EC 6.1.1.5).
 GN Names-iles; OrderedlocusNames=BT9727.3637;
 OS Bacillus thuringiensis (subsp. konkukian).
 OC Bacteria; Firmicutes; Bacilliales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=180856;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=97-27;
 RA Brettin T.S., Bruce D., Challacombe J.F., Gilna P., Han C., Hill K., Hitchcock P., Jackson P., Keim P., Longmire J., Lucas S., Okinaka R., Richardson P., Rubin E., Tice H.;
 RA "Complete genome sequence of Bacillus thuringiensis 97-27.";
 RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AE017355; AAT63892.1; -.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0004822; P:isoleucine-tRNA ligase activity; IEA.
 DR GO; GO:0006428; P:isoleucyl-tRNA aminoacylation; IEA.
 DR InterPro; IPR002300; tRNA-synt la.
 DR InterPro; IPR001412; tRNA-synt-1.
 DR InterPro; IPR002301; tRNA-synt-ile.
 DR InterPro; IPR009080; tRNA-synt-ile.
 DR InterPro; IPR009080; tRNA-synt-ile.
 DR Pfam; PF001133; tRNA-synt 1; 1.
 DR Pfam; PF06827; zf-PFG ilers; 1.
 DR PRINTS; PR00984; tRNA-synt-1.
 DR TIGRFAMs; TIGR00392; iles; 1.
 DR PROSITE; PS00178; AA_TRNA_LIGASE I; 1.
 KW Complete proteome.
 SQ SEQUENCE 921 AA; 104633 MW; ABF082F44E94A891 CRC64;

Query Match 82.8%; Score 24; DB 2; Length 921;
 Best Local Similarity 83.3%; Pred. No. 1.6e+03;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 MKVAEP 7
 | | | | |
 Db 115 MSVAEP 120

RESULT 82

Q65JV7 PRELIMINARY; PRT; 922 AA.
 ID Q65JV7
 AC Q65JV7
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
 DE Iles [Isoleucyl-tRNA synthetase].
 GN Names-iles; ORFNames=BL02267, BL101762;
 OS Bacillus licheniformis DSM 13.
 OC Bacteria; Firmicutes; Bacilliales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=279010;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=DSM 13;
 RX PubMed=15383718;
 RA Veith B., Herzberg C., Streckel S., Feesche J., Maurer K.H., Ehrenreich P., Baumer S., Henne A., Liesegang H., Merkl R., Ehrenreich A., Gottschalk G.;
 RA "The Complete Genome Sequence of Bacillus licheniformis DSM13, an Organism with Great Industrial Potential.";
 RT J. Mol. Microbiol. Biotechnol. 7:204-211(2004).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 14580;
 RA Rey M.W., Ramaiya P., Nelson B.A., Brody-Karpin S.D., Zaretsky E.J., Tang M., de Leon A.L., Xiang H., Gusti V., Clausen I.G., Olsen P.B., Rasmussen M.D., Andersen J.T., Jorgensen P.L., Larsen T.S., Sorokin A., Bolotin A., Lapidus A., Galleron N., Ehrlich S.D., Berka R.M.;
 RA "Complete genome sequence of the industrial bacterium Bacillus licheniformis and comparisons with closely related Bacillus species.";
 RL Genome Biol. 5:R77-R77(2004).
 DR EMBL; AE017333; AAU40657.1; -.
 DR EMBL; CP000002; AAU23298.1; -.
 KW Aminoacyl-tRNA synthetase.
 SQ SEQUENCE 922 AA; 104355 MW; DFCB33B0F71906B6 CRC64;

Query Match 82.8%; Score 24; DB 2; Length 922;
 Best Local Similarity 83.3%; Pred. No. 1.6e+03;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 MKVAEP 7
 | | | | |
 Db 116 MTVAEP 121

RESULT 83

Q6G4C0 PRELIMINARY; PRT; 961 AA.
 ID Q6G4C0
 AC Q6G4C0
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Hypothetical protein.
 GN OrderedlocusNames=BH04460;
 OS Bartonella henselae (Rochalimaea henselae).
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales; Bartonellaceae; Bartonella.
 OX NCBI_TaxID=38323;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 49882 / Houston 1;
 RX PubMed=15210978; DOI=10.1073/pnas.0305659101;
 RA Alsmark U.C.M., Frank A.C., Karlberg E.O., Legault B.-A., Ardell D.H., Canbaeck B., Eriksson A.-S., Naeslund A.K., Handley S.A., Huvet M., La Scola B., Holmberg M., Andersson S.G.E.;
 RA "The louse-borne human pathogen Bartonella quintana is a genomic derivative of the zoonotic agent Bartonella henselae.";
 RL Proc. Natl. Acad. Sci. U.S.A. 101:9716-9721(2004).
 DR EMBL; BX897699; CAF27255.1; -.

Query Match 82.8%; Score 24; DB 2; Length 921;
 Best Local Similarity 83.3%; Pred. No. 1.6e+03;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 MKVAEP 7
 | | | | |
 Db 115 MSVAEP 120

GO; GO:0000155; F:two-component sensor molecule activity; IEA.
GO; GO:0000160; P:two-component signal transduction system (p. . .; IEA.
InterPro; IPR001633; EAL.
InterPro; IPR000160; GDEF.
InterPro; IPR000014; PAS.
InterPro; IPR000700; PAS-assoc_C.
PFam; PF00563; EAL; 1.
PFam; PF00990; GDEF; 1.
SMART; SM00267; DUF1; 1.
SMART; SM00052; DUF2; 1.
TIGRFAMs; TIGR00254; GDEF; 1.
TIGRFAMs; TIGR00229; sensory_box; 1.
PROSITE; PS00883; EAL; 1.
PROSITE; PS00887; GDEF; 1.
PROSITE; PS01113; PAC; 1.
Complete proteome; Hypothetical protein.
KW COMPLETE PROTEOME; Hypothetical protein.
SQ SEQUENCE 961 AA; 108536 MW; BA29615773580526 CRC64;
Query Match 82.8%; Score 24; DB 2; Length 961;
Best Local Similarity 71.4%; Pred. No. 1.6e+03;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 VMXVAEF 7
DB 765 VMDLAEF 771
RESULT 84
AMPN_HAECO STANDARD; PRT; 971 AA.
ID AMPN_HAECO STANDARD; PRT; 971 AA.
AC Q10737;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Aminopeptidase N (EC 3.4.11.2) (Microsomal aminopeptidase) (Membrane
DE Glycoprotein Hil).
OS Haemochus contortus (Barber pole worm).
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;
OC Trichostrongyloidea; Haemonchidae; Haemonchinae; Haemonchus.
OX NCBI_TaxID=6289;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97273974; PubMed=9128148; DOI=10.1016/S0167-4838(96)00204-X;
RA Smith T.S., Graham M., Munn E.A., Newton S.E., Knox D.P.,
RA Coadwell W.J., McMichael-Phillips D., Smith H., Smith W.D.,
RA Oliver J.J.;
RT "Cloning and characterization of a microsomal aminopeptidase from the
RT intestine of the nematode Haemochus contortus.";
RL Biochim. Biophys. Acta 1338:295-306(1997).
CC -1- CATALYTIC ACTIVITY: Release of an N-terminal amino acid, Xaa-|-
CC Xbb- from a peptide, amide or arylamide. Xaa is preferably Ala,
CC but may be most amino acids including Pro (slow action). When a
CC terminal hydrophobic residue is followed by a prolyl residue, the
CC two may be released as an intact Xaa-Pro dipeptide.
CC -1- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
CC -1- SUBCELLULAR LOCATION: Type II membrane protein (By similarity).
CC -1- SIMILARITY: Belongs to the peptidase M1 family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X94187; CAA63897.1; --
DR MSERP; M01.015; --
DR InterPro; IPR006025; Pept_M_Zn_BS.
DR InterPro; IPR001930; Peptidase_M1.
DR Pfam; PF01433; Peptidase_M1; 1.
DR PRINTS; PR00756; ALADIPTASE.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.

KW Aminopeptidase; Glycoprotein; Hydrolase; Metalloprotease;
KW Signal-anchor; Transmembrane; Zinc.
FT INIT_MET 0 By similarity.
FT DOMAIN 1 16 Cytoplasmic (Potential).
FT TRANSMEM 17 38 Signal-anchor for type II membrane
FT protein (Potential).
FT DOMAIN 39 971 Extracellular (Potential).
FT METAL 378 378 Zinc (catalytic) (By similarity).
FT ACT_SITE 379 379 By similarity.
FT METAL 382 382 Zinc (catalytic) (By similarity).
FT METAL 401 401 Zinc (catalytic) (By similarity).
FT ACT_SITE 465 465 Proton donor (Potential).
FT CARBOHYD 98 98 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 226 226 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 548 548 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 857 857 N-linked (GlcNAc...) (Potential).
SQ SEQUENCE 971 AA; 110542 MW; 95C6A92B5CCA227C CRC64;
Query Match 82.8%; Score 24; DB 1; Length 971;
Best Local Similarity 71.4%; Pred. No. 1.6e+03;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 VMXVAEF 7
DB 541 VISVAEF 547
RESULT 85
Q86G73 PRELIMINARY; PRT; 972 AA.
ID Q86G73 PRELIMINARY; PRT; 972 AA.
AC Q86G73;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hidden antigen h11.
OS Haemochus contortus (Barber pole worm).
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;
OC Trichostrongyloidea; Haemonchidae; Haemonchinae; Haemonchus.
OX NCBI_TaxID=6289;
RN [1]
RP SEQUENCE FROM N.A.
RA Yan R.F., Li X.R.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY247714; AAO91936.1; --
DR GO; GO:0004179; F:membrane alanyl aminopeptidase activity; IEA.
DR GO; GO:0005508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001930; Peptidase_M1.
DR InterPro; IPR006025; Pept_M_Zn_BS.
DR Pfam; PF01433; Peptidase_M1; 1.
DR PRINTS; PR00756; ALADIPTASE.
DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN 1.
SQ SEQUENCE 972 AA; 110548 MW; 8DBC651B21D905D4 CRC64;
Query Match 82.8%; Score 24; DB 2; Length 972;
Best Local Similarity 71.4%; Pred. No. 1.6e+03;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 VMXVAEF 7
DB 542 VISVAEF 548
RESULT 86
AT9B_HUMAN STANDARD; PRT; 1095 AA.
ID AT9B_HUMAN STANDARD; PRT; 1095 AA.
AC O43861; O60872;
DT 30-MAY-2000 (Rel. 39, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Potential phospholipid-transporting ATPase 11B (EC 3.6.3.1) (HUSSY-
DE 20).
GN Name=ATP9B; Synonyms=ATPIIB, NEOL1;
OS Homo sapiens (Human).

OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	RC	TISSUE=Brain;	
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	RX	MEDLINE=21064499; PubMed=11124703;	
OX	NCBI_TaxID=9606;	RX	DOI=10.1002/1097-0061(200101)18:1<69::AID-YEA647>3.3.CO;2-8;	
RN	[1]	RA	Stanchi F., Bertocco E., Toppo S., Dioguardi R., Simonati B.,	
RP	SEQUENCE OF 1-420 FROM N.A.	RA	Cannata N., Zimbello R., Lanfranchi G., Valle G.;	
RC	TISSUE=Testis;	RT	"Characterization of 16 novel human genes showing high similarity to	
RA	PubMed=14702039; DOI=10.1038/ng1285;	RL	Yeast sequences.";	
RA	Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,	RN	Yeast 18:69-80(2001).	
RA	Wakanatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,	RP	[5]	
RA	Sekine M., Obayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,	RA	Axelsen K.B.;	
RA	Yamanoto J.-I., Saito K., Kawai Y., Isono Y., Nakamura Y.,	RL	Unpublished observations (JUL-2002).	
RA	Nagahori K., Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M.,	CC	!- CATALYTIC ACTIVITY: ATP + H(2)O = ADP + phosphate.	
RA	Shiratori A., Sudo H., Hosoi T., Kaku Y., Kodaira H., Kondo H.,	CC	!- SUBCELLULAR LOCATION: Integral membrane protein.	
RA	Sugawara M., Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E.,	CC	!- SIMILARITY: Belongs to the cation transport ATPases family (P-type	
RA	Yamazaki M., Ninomiya K., Ishibashi T., Yamashita H., Murakawa K.,	CC	ATPases). Subfamily IV.	
RA	Fujimori K., Tanai H., Kimata M., Watanabe M., Hiraoka S., Chiba Y.,	CC	-----	
RA	Ishida S., Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T.,	CC	This SWISS-PROT entry is copyright. It is produced through a collaboration	
RA	Kusano J., Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O.,	CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -	
RA	Nomura Y., Togiya S., Konai F., Hara R., Takeuchi K., Arita M.,	CC	the European Bioinformatics Institute. There are no restrictions on its	
RA	Imose N., Musashino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,	CC	use by non-profit institutions as long as its content is in no way	
RA	Yoshihaka Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,	CC	modified and this statement is not removed. Usage by and for commercial	
RA	Moriya S., Momiya H., Satoh N., Takami S., Terashima Y., Suzuki O.,	CC	entities requires a license agreement (See http://www.isb-sib.ch/	
RA	Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,	CC	or send an email to license@isb-sib.ch).	
RA	Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,	CC	-----	
RA	Fujimori Y., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,	DR	EMBL; AK097757; -; NOT ANNOTATED_CDS.	
RA	Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y.,	DR	EMBL; AC104423; -; NOT ANNOTATED_CDS.	
RA	Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,	DR	EMBL; U78978; AAC05243.1; -.	
RA	Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,	DR	EMBL; AJ006268; CAA06934.1; -.	
RA	Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,	DR	Genew; HGNC:13541; ATP9B.	
RA	Togashi T., Oyama M., Hara H., Watanabe M., Komatsu T.,	DR	GO; GO:0016021; C:integral to membrane; NAS.	
RA	Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,	DR	GO; GO:0015247; F:aminophospholipid transporter activity; NAS.	
RA	Okumura K., Nagase T., Nomura N., Kikuchi H., Masuo Y., Yamashita R.,	DR	GO; GO:0005524; F:ATP binding; NAS.	
RA	Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;	DR	GO; GO:0019829; F:cation-transporting ATPase activity; NAS.	
RT	"Complete sequencing and characterization of 21,243 full-length human	DR	GO; GO:0015917; P:aminophospholipid transport; NAS.	
RT	cDNAs.";	DR	InterPro; IPR001757; Dehal_Like_hydro.	
RL	Nat. Genet. 36:40-45(2004).	DR	InterPro; IPR005834; Dehal_Like_hydro.	
RN	[2]	DR	InterPro; IPR008250; E1-E2_ATPase_reg.	
RP	SEQUENCE OF 208-1095 FROM N.A.	DR	InterPro; IPR006539; Flippase.	
RA	Birren B., Linton L., Nubstien C., Lander E., Ali A., Allen N.,	DR	Pfam; PF00122; E1-E2_ATPase; 1.	
RA	Anderson S., Barna N., Bastien V., Boguslavsky L., Boukhgalter B.,	DR	Pfam; PF00702; Hydrolase; 1.	
RA	Brown A., Camarata J., Campopiano A., Chang J., Chazaro B.,	DR	PRINTS; PR00119; CATATPASE.	
RA	Choepel Y., Collangelo M., Collins S., Collymore A., Cook A., Cooke P.,	DR	TIGRFAMs; TIGR01652; ATPase-Plipid; 1.	
RA	Deaellano K., Dewar K., Diaz J.S., Dodge S., Faro S., Ferreira P.,	DR	PROSITE; PS00154; ATPASE_E1_E2; 1.	
RA	Fitzhugh W., Gage D., Galagan J., Gardyna S., Ginde S., Gord S.,	KW	ATP-binding; Hydrolase; Magnesium; Multigene family; Phosphorylation;	
RA	Goyette M., Graham L., Grand-Pierre N., Hagos B., Heaford A.,	KW	Transmembrane.	
RA	Horton L., Hulme W., Iliev I., Johnson R., Jones C., Kamat A.,	FT	DOMAIN 1 94	
RA	Karatas A., Kells C., LaRoque K., Lamazares R., Landers T.,	FT	TRANSMEM 95 116	Cytoplasmic (Potential).
RA	Lehoczyk J., Levine R., Liu G., MacLean C., Macdonald P., Major J.,	FT	DOMAIN 117 121	Extracellular (Potential).
RA	Marquis N., Matthews C., McCarthy M., McEwan P., McKernan K.,	FT	TRANSMEM 122 144	Potential.
RA	McPeeters R., Meldrum J., Meneus L., Mihova T., Mienga V., Murphy T.,	FT	DOMAIN 145 328	Cytoplasmic (Potential).
RA	Naylor J., Nguyen C., Norbu C., Norman C.H., O'Connor T.,	FT	TRANSMEM 329 349	Potential.
RA	O'Donnell P., O'Neill D., Oliver J., Peterson K., Phunkhang P.,	FT	DOMAIN 350 357	Extracellular (Potential).
RA	Pierre N., Pollara V., Raymond C., Retta R., Rieback M., Riley R.,	FT	TRANSMEM 358 379	Potential.
RA	Rise C., Rogov P., Roman J., Rosetti M., Roy A., Santos R.,	FT	DOMAIN 380 878	Cytoplasmic (Potential).
RA	Schauer S., Schupack R., Seaman S., Severy P., Spencer B.,	FT	TRANSMEM 879 899	Potential.
RA	Stange-Thomann N., Stojanovic N., Strauss N., Subramanian A.,	FT	DOMAIN 900 911	Extracellular (Potential).
RA	Talanas J., Tesfaye S., Theodore J., Topham K., Travers M., Travis N.,	FT	TRANSMEM 912 930	Potential.
RA	Trigilio J., Vassiliev H., Viel R., Vo A., Wilson B., Wu X., Wyman D.,	FT	DOMAIN 931 960	Cytoplasmic (Potential).
RA	Ye W.J., Young G., Zainov J., Zembek L., Zimmer A., Zody M.;	FT	TRANSMEM 961 979	Potential.
RL	Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.	FT	DOMAIN 980 986	Extracellular (Potential).
RN	[3]	FT	TRANSMEM 987 1009	Potential.
RP	SEQUENCE OF 759-1095 FROM N.A.	FT	DOMAIN 1010 1015	Cytoplasmic (Potential).
RC	TISSUE=Brain;	FT	TRANSMEM 1016 1036	Potential.
RX	MEDLINE=98217376; PubMed=9548971;	FT	DOMAIN 1037 1053	Extracellular (Potential).
RA	Halleck M.S., Pradhan D., Blackman C.F., Berkes C., Williamson P.L.,	FT	TRANSMEM 1054 1078	Potential.
RA	Schlegel R.A.;	FT	DOMAIN 1079 1095	Cytoplasmic (Potential).
RT	"Multiple members of a third subfamily of P-type ATPases identified by	FT	ACT_SITE 416 416	4-aspartylphosphate intermediate (By
RT	genomic sequences and ESTs.";	FT	METAL 822 822	similarity).
RL	Genome Res. 8:354-361(1998).	FT	METAL 826 826	Magnesium (By similarity).
RN	[4]	FT	CONFLICT 834 834	Magnesium (By similarity).
RP	SEQUENCE OF 768-1095 FROM N.A.	FT	CONFLICT 954 954	D -> N (in Ref. 3).
				R -> K (in Ref. 3).

```

FT CONFLICT 968 968 S -> I (in Ref. 3).
FT CONFLICT 984 984 E -> D (in Ref. 3).
FT CONFLICT 1021 1021 E -> D (in Ref. 3).
FT CONFLICT 1043 1043 R -> K (in Ref. 3).
FT CONFLICT 1051 1051 D -> N (in Ref. 3).
FT CONFLICT 1079 1079 K -> N (in Ref. 3).
SQ SEQUENCE 1095 AA; 123635 MW; 570C2B7D8507C255 CRC64;

Query Match 82.8%; Score 24; DB 1; Length 1095;
Best Local Similarity 71.4%; Pred. No. 1.8e+03;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
Db 1016 LMVVAEF 1022

RESULT 87
AT9B_MOUSE STANDARD; PRT; 1095 AA.
AC P98135; Q99LI3;
DT 30-MAY-2000 (Rel. 39, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE 05-JUN-2004 (Rel. 44, Last annotation update)
DE Potential phospholipid-transporting ATPase IIB (BC 3.6.3.1).
GN Name=At9B;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=ICR; TISSUE=Brain;
RX MEDLINE=20473714; PubMed=11015572;
RA Halleck M.S., Lawler J.F. Jr., Blackshaw S., Gao L., Nagarajan P.,
RA Hacker C., Fyle S., Newman J.T., Nakanishi Y., Ando H., Weinstein D.,
RA Williamson P., Schlegel R.A.;
RT "Differential expression of putative transbilayer amphipath
RT transporters.";
RL Physiol. Genomics 1:139-150(1999).
RN [2]
SEQUENCE OF 658-1095 FROM N.A.
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -I- CATALYTIC ACTIVITY: ATP + H(2)O = ADP + phosphate.
CC -I- SUBCELLULAR LOCATION: Integral membrane protein.
CC -I- TISSUE SPECIFICITY: Found in most tissues except spleen and
CC muscle. Most abundant in testis. Also detected in fetal tissues.
CC -I- SIMILARITY: Belongs to the cation transport ATPases family (P-type
CC ATPases). Subfamily IV.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```

```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF155913; AAF08476.1; -.
DR EMBL; BC003246; AAO03246.1; -.
DR MGD; MGI:1354757; At9B.
DR InterPro; IPR001757; ATPase_E1-E2.
DR InterPro; IPR005834; Dehal_like_hydro.
DR InterPro; IPR008250; E1-E2_ATPase_reg.
DR InterPro; IPR006539; Flippase.
DR Pfam; PF00122; E1-E2_ATPase; 1.
DR Pfam; PF00702; Hydrolase; 1.
DR PRINTS; PR00119; CATATPASE.
DR TIGRFAMs; TIGR01652; ATPase-Plipid; 1.
DR TIGRFAMs; TIGR01494; ATPase_P-type; 7.
DR PROSITE; PS00154; ATPASE_E1-E2; 1.
DR KW ATP-binding; Hydrolase; Magnesium; Multigene family; Phosphorylation;
KW Transmembrane.
FT DOMAIN 1 94 Cytoplasmic (Potential).
FT TRANSMEM 95 116 Potential.
FT DOMAIN 117 121 Extracellular (Potential).
FT TRANSMEM 122 144 Potential.
FT DOMAIN 145 328 Cytoplasmic (Potential).
FT TRANSMEM 329 349 Potential.
FT DOMAIN 350 357 Extracellular (Potential).
FT TRANSMEM 358 379 Potential.
FT DOMAIN 380 878 Cytoplasmic (Potential).
FT TRANSMEM 879 899 Potential.
FT DOMAIN 900 911 Extracellular (Potential).
FT TRANSMEM 912 930 Potential.
FT DOMAIN 931 960 Cytoplasmic (Potential).
FT TRANSMEM 961 979 Potential.
FT DOMAIN 980 986 Extracellular (Potential).
FT TRANSMEM 987 1009 Potential.
FT DOMAIN 1010 1015 Cytoplasmic (Potential).
FT TRANSMEM 1016 1036 Potential.
FT DOMAIN 1037 1053 Extracellular (Potential).
FT TRANSMEM 1054 1078 Potential.
FT DOMAIN 1079 1095 Cytoplasmic (Potential).
FT ACT_SITE 416 416 4-aspartylphosphate intermediate (By
similarity).
FT METAL 822 822 Magnesium (By similarity).
FT METAL 826 826 Magnesium (By similarity).
FT CONFLICT 1040 1050 Missing (in Ref. 2).
SQ SEQUENCE 1095 AA; 123546 MW; A0176FEASDFEAL79 CRC64;

Query Match 82.8%; Score 24; DB 1; Length 1095;
Best Local Similarity 71.4%; Pred. No. 1.8e+03;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
Db 1016 LMVVAEF 1022

RESULT 88
VABI_CABEL STANDARD; PRT; 1122 AA.
ID VABI_CABEL
AC 061460; Q21477;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Ephrin receptor 1 precursor (EC 2.7.1.112) (Tyrosine-protein kinase
DE Names=vab-1; ORFNames=W03A1.1;
GN Caenorhabditis elegans.
OC Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
SEQUENCE FROM N.A.; FUNCTION, MUTANTS, AND TISSUE SPECIFICITY.

```


RC STRAIN=Bristol N2;
RX MEDLINE=98165343; PubMed=9506518; DOI=10.1016/S0092-8674(00)81131-9;
RA George S.E., Smoket K., Haroin J., Chisholm A.D.;
RT "The VAB-1 Eph receptor tyrosine kinase functions in neural and
RT epithelial morphogenesis in *C. elegans*.";
RL Cell 92:633-643(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RG The *C. elegans* sequencing consortium;
RT "Genome sequence of the nematode *C. elegans*: a platform for
RT investigating biology.";
RL Science 282:2012-2018(1998).
RN [3]
RP REVISIONS, AND ALTERNATIVE SPLICING.
RA Waterston R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Receptor for members of the ephrin family. Involved in
CC interactions between neuronal substrate cells and a migrating
CC epithelial sheet in head epidermis morphogenesis. Also required
CC for cell movements following gastrulation and during ventral
CC closure of the epidermis.
CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=a;
CC IsoId=O61460-1; Sequence=Displayed;
CC Name=b;
CC IsoId=O61460-2; Sequence=VSP_050207;
CC Note=No experimental confirmation available;
CC -1- TISSUE SPECIFICITY: During ventral enclosure of the epidermis,
CC expression is seen in clusters of presumptive head neuronal cells
CC and several cells in the tail region. Early larvae show expression
CC in the nerve ring and ventral nerve cord. Strong expression is
CC also seen in the procorpus and terminal bulb of the pharynx.
CC Expression in the nervous system is seen through to adulthood.
CC -1- SIMILARITY: Belongs to the Tyr protein kinase family. Ephrin
CC receptor subfamily.
CC -1- SIMILARITY: Contains 2 fibronectin type III domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF040269; AAC38970.1; -;
CC EMBL; U49956; AAK77620.1; -;
CC EMBL; U49956; AAK77621.1; -;
CC PIR; T42400; T42400.
CC HSP; P54763; IJPA.
CC WormBase; WBGene00006868; vab-1.
CC WormPep; M03A1.1a; CE25060.
CC WormPep; M03A1.1b; CE28617.
CC GO; GO:0016021; C:integral to membrane; ISS.
CC GO; GO:0005003; F:ephrin receptor activity; ISS.
CC GO; GO:0004713; F:protein-tyrosine kinase activity; ISS.
CC GO; GO:0007391; F:dorsal closure; IMP.
CC InterPro; IPR001090; Ephrin_receptor.
CC InterPro; IPR003961; FN III.
CC InterPro; IPR008957; FN_III-like.
CC InterPro; IPR008979; Gal_bind_like.
CC InterPro; IPR011009; Kinase_like.
CC InterPro; IPR000719; Prot_kinase.
CC InterPro; IPR001245; Tyr_kinase.
CC InterPro; IPR008266; Tyr_pkinase_AS.
CC Pfam; PF01404; Ephrin_lbd; 1.
CC Pfam; PF00041; fn3; 2.

DR Pfam; PF00069; Pkinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD001495; Ephrin_receptor; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00615; EPH_lbd; 1.
DR SMART; SM00060; FN3; 2.
DR SMART; SM00219; TyrKc; 1.
DR PROSITE; PS0853; FN3; 2.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS0011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; FALSE NEG.
KW Alternative splicing; ATP-binding; Developmental protein;
KW Glycoprotein; Phosphorylation; Receptor; Repeat; Signal; Transferase;
KW Transmembrane; Tyrosine-protein kinase.
FT SIGNAL 1 29 Potential.
FT CHAIN 30 1122 Ephrin receptor 1.
FT DOMAIN 30 559 Extracellular (Potential).
FT TRANSMEM 560 580 Potential.
FT DOMAIN 581 1122 Cytoplasmic (Potential).
FT DOMAIN 199 338 Cys-rich.
FT DOMAIN 341 436 Fibronectin type-III 1.
FT DOMAIN 450 536 Fibronectin type-III 2.
FT DOMAIN 692 988 Protein kinase.
FT SITE 1120 1122 PDZ-binding motif (Potential).
FT NP_BIND 698 706 ATP (By similarity).
FT BINDING 748 748 ATP (By similarity).
FT ACT_SITE 849 849 By similarity.
FT MOD_RES 667 667 Phosphotyrosine (by autocatalysis)
FT MOD_RES 673 673 Phosphotyrosine (by autocatalysis)
FT MOD_RES 885 885 Phosphotyrosine (by autocatalysis)
FT MOD_RES 1048 1048 Phosphotyrosine (by autocatalysis)
FT CARBOHYD 172 172 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 238 238 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 244 244 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 499 499 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 554 554 N-linked (GlcNAc...) (Potential).
FT VARSPLIC 843 847 Missing (in isoform b).
FT FTId=VSP_050207.
FT MUTAGEN 62 62 E->K: In allele ju8; most die as larvae.
FT MUTAGEN 63 63 T->I: In allele e699; most die as larvae.
FT MUTAGEN 195 195 E->K: In allele e856; most die as larvae.
FT MUTAGEN 917 917 G->E: In allele e2; most die as adults.
FT MUTAGEN 966 966 C->F: In allele ju22; most die as adults.
FT MUTAGEN 966 966 C->Y: In allele e1063; most die as adults.
SQ SEQUENCE 1122 AA; 125263 MW; 6E4A3037BB92A1D0 CRC64;
Query Match 82.8%; Score 24; DB 1; Length 1122;
Best Local Similarity 71.4%; Pred. No. 1.9e+03;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 VMXVAEF 7
DB 790 VMVAEY 796
RESULT 89
Q965S4 PRELIMINARY; PRT; 1130 AA.
AC Q965S4
ID Q965S4;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-WAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein Y75B7AL.4.
GN Name=Y75B7AL.4; ORFNames=Y75B7AL.4;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;

RN [1] SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RG MEDLINE=99069613; PubMed=9851916;
 RX WormBase Consortium;
 RT "Genome sequence of the nematode *C. elegans*: a platform for
 RL investigating biology. The *C. elegans* Sequencing Consortium.";
 RL Science 282:2012-2018(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Lamar B.;
 RT "The sequence of *C. elegans* cosmid Y75B7AL.";
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Waterston R.H.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Waterston R.;
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Waterston R.;
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Waterston R.;
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
 RN [7]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Waterston R.;
 RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
 RN [8]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Waterston R.;
 RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
 RN [9]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Waterston R.;
 RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
 RN [10]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RG WormBase Consortium;
 RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC024869; AAK68606.2; -;
 DR WormBase; WBGene00022286; Y75B7AL.4.
 DR WormPep; Y75B7AL.4; CE32278.
 DR InterPro; IPR000198; RhoGAP.
 DR InterPro; IPR008936; Rho GAP.
 DR Pfam; PF00620; RhoGAP; 1.
 DR SMART; SM00324; RhoGAP; 1.
 DR PROSITE; PS50238; RHO GAP; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 1130 AA; 123777 MW; FAC055A8A919C8E4 CRC64;

Query Match 82.8%; Score 24; DB 2; Length 1130;
 Best Local Similarity 83.3%; Pred. No. 1.9e+03;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 2 MVVAEF 7
 Db 537 MSVAEF 542

RESULT 90
 Q68FM3 PRELIMINARY; PRT; 1146 AA.
 ID O68FM3
 AC O68FM3
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
 DE Atp9b protein.
 GN Name=Atp9b;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Brain;
 RX PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Vallalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RL and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Brain;
 RA Director MGC Project;
 RL Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC079626; AAH79626.1; -;
 DR InterPro; IPR001757; ATPase E1-E2.
 DR InterPro; IPR005834; Dehal-like_hydro.
 DR InterPro; IPR006539; Flippase.
 DR Pfam; PF00702; Hydrolase; 1.
 DR PRINTS; PR00119; CATATPASE.
 DR TIGRFAMs; TIGR01652; ATPase-Plipid; 1.
 DR TIGRFAMs; TIGR01494; ATPase P-type; 4.
 DR PROSITE; PS00154; ATPASE_E1_E2; UNKNOWN 1.
 SQ SEQUENCE 1146 AA; 129044 MW; FF07AFD694A1830E CRC64;
 Query Match 82.8%; Score 24; DB 2; Length 1146;
 Best Local Similarity 71.4%; Pred. No. 1.9e+03;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 VMXVAEF 7
 Db 1067 LMVVAEF 1073
 RESULT 91
 Q868R0 PRELIMINARY; PRT; 1173 AA.
 ID Q868R0
 AC Q868R0;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Reverse transcriptase.
 OS Anopheles gambiae (African malaria mosquito).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

```

OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.
OX NCBI_TaxID=7165;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22531580; PubMed=12644555; DOI=10.1093/molbev/msg031;
RA Kojima K.K., Fujiwara H.;
RT "Evolution of target specificity in R1 clade non-LTR
retrotransposons";
RL Mol. Biol. Evol. 20:351-361(2003).
DR EMBL; AB090822; BAC57920.1; -.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006278; P:RNA-dependent DNA replication; IEA.
DR InterPro; IPR005135; Exo_endo_phos.
DR InterPro; IPR002088; PPTA.
DR InterPro; IPR000477; RVtse.
DR Pfam; PF00372; Exo_endo_phos; 1.
DR PROSITE; PS00904; PPTA; UNKNOWN_1.
KW RNA-directed DNA polymerase; Transferase.
SQ SEQUENCE 1173 AA; 132166 MW; AF60298E9A9D67B5 CRC64;

Query Match 82.8%; Score 24; DB 2; Length 1173;
Best Local Similarity 83.3%; Pred. No. 2e+03;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 MVXAEF 7
DB 94 MTVAEF 99

RESULT 92
Q7NEB4
ID Q7NEB4 PRELIMINARY; PRT; 1566 AA.
AC Q7NEB4;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Gll3966 protein.
GN OrderedLocusNames=gll3966;
OS Gloeobacter violaceus.
OC Bacteria; Cyanobacteria; Chroococcales; Gloeobacter.
OX NCBI_TaxID=33072;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PCC 7421;
RX MEDLINE=22977040; PubMed=14621292;
RA Nakamura Y., Kaneko T., Sato S., Mimuro M., Miyashita H., Tsuchiya T.,
RA Sasanoto S., Watanabe A., Kawashima K., Kishida Y., Kiyokawa C.,
RA Kohara M., Matsumoto M., Matsuno A., Nakazaki N., Shimpo S.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of Gloeobacter violaceus PCC 7421, a
cyanobacterium that lacks thylakoids.";
RL DNA Res. 10:137-145(2003).
DR EMBL; AP006581; BAC91907.1; -.
DR HSSP; Q55418; 1W18.
DR GO; GO:0004519; F:endonuclease activity; IEA.
DR GO; GO:0016539; P:intron-mediated protein splicing; IEA.
DR InterPro; IPR006142; INTEIN.
DR InterPro; IPR004042; Intein endonuc.
DR InterPro; IPR006141; Intein_S.
DR PRINTS; PR00379; INTEIN.
DR TIGRFAMs; TIGR01443; Intein_Cterm; 2.
DR TIGRFAMs; TIGR01445; Intein_Nterm; 1.
DR PROSITE; PS50818; INTEIN_C_TER; 2.
DR PROSITE; PS50819; INTEIN_ENDONUCLEASE; 2.
DR PROSITE; PS50817; INTEIN_N_TER; 2.
KW Complete proteome.
SQ SEQUENCE 1566 AA; 172226 MW; 907814D1219662A5 CRC64;

Query Match 82.8%; Score 24; DB 2; Length 1566;
Best Local Similarity 71.4%; Pred. No. 2.6e+03;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
DB 1173 VLDVAEF 1179

RESULT 93
Q7XI95
ID Q7XI95 PRELIMINARY; PRT; 2033 AA.
AC Q7XI95;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein P0453G03.24.
GN Names=P0453G03.24;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 7, PAC
clone:P0453G03.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP004276; BAC79823.1; -.
DR Gramene; Q7XI95; -.
DR InterPro; IPR000413; Integrin alpha.
DR PROSITE; PS00242; INTEGRIN_ALPHA; UNKNOWN_1.
KW Hypothetical protein..
SQ SEQUENCE 2033 AA; 221848 MW; 247BFD68BD43562C CRC64;

Query Match 82.8%; Score 24; DB 2; Length 2033;
Best Local Similarity 57.1%; Pred. No. 3.3e+03;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
DB 1843 ILAVAEF 1849

RESULT 94
Q70KH4
ID Q70KH4 PRELIMINARY; PRT; 2201 AA.
AC Q70KH4;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Polyketide synthase type I.
GN Names=aurC;
OS Streptomyces thioluteus.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=66431;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HKI-227;
RX PubMed=14700630; DOI=10.1016/j.chembiol.2003.11.009;
RA He J., Hertweck C.;
RT "Iteration as programmed event during polyketide assembly; molecular
analysis of the aureothin biosynthesis gene cluster.";
RL Chem. Biol. 10:1225-1232(2003).
DR EMBL; AJ575648; CAE02606.1; -.
DR HSSP; Q03133; 1KEZ.
DR GO; GO:0048037; F:cofactor binding; IEA.
DR GO; GO:0016788; F:hydrolyase activity; acting on ester bonds; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0009058; P:biosynthesis; IEA.
DR GO; GO:0006633; P:fatty acid biosynthesis; IEA.
DR InterPro; IPR009081; ACP like.
DR InterPro; IPR001227; Ac_transferase.

```

```

DR InterPro; IPR000794; Ketoacyl_synth.
DR InterPro; IPR008262; Lipase AS.
DR InterPro; IPR006163; Phosphatidyleth_bind.
DR InterPro; IPR006162; Phantne S..
DR InterPro; IPR001031; Thioesterase.
DR Pfam; PF00698; Acyl_transf_1; 2.
DR Pfam; PF00109; Ketoacyl-synt; 2.
DR Pfam; PF02801; Ketoacyl-synt_C; 2.
DR Pfam; PF00550; pp-binding; 2.
DR Pfam; PF00975; Thioesterase; 1.
DR PROSITE; PS50075; ACP DOMAIN; 2.
DR PROSITE; PS00606; E_KETOACYL_SYNTHASE; 2.
DR PROSITE; PS00120; LIPASE SER; UNKNOWN; 1.
DR PROSITE; PS00012; PHOSPHOPANTETHEINE; 1.
DR PHOSPHOPANTHEINE; TRANSFERASE.
SQ SEQUENCE 2201 AA; 233522 MW; 11440F4B7192A336 CRC64;

Query Match 82.8%; Score 24; DB 2; Length 2201;
Best Local Similarity 71.4%; Pred. No. 3.6e+03;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VMXVAEF 7
|: ||||
Db 1954 VLRVAEF 1960

RESULT 95
Q712V2 PRELIMINARY; PRT; 53 AA.
AC Q712V2;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Protein tyrosine kinase (EC 2.7.1.12) (Fragment).
GN Name=DmHD-311;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA Oates A.C., Wollberg P., Achen M.G., Wilks A.F.;
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ002918; CAA05753.1; -.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR000719; Prot_kinase.
DR ProDom; PD000001; Prot_kinase; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW Kinase; Transferase.
FT NON_TER 1
FT NON_TER 53
SQ SEQUENCE 53 AA; 6170 MW; 5AB107970ECA34F7 CRC64;

Query Match 79.3%; Score 23; DB 2; Length 53;
Best Local Similarity 57.1%; Pred. No. 2e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VMXVAEF 7
|: ||||
Db 8 VMKIADF 14

RESULT 96
Q91995 PRELIMINARY; PRT; 56 AA.
AC Q91995;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)

```

```

DE G62 protein (E7 protein) (Fragment).
GN Name=G62;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=95383727; PubMed=7655077;
RA Brandli A.W., Kirschner M.W.;
RT "Molecular cloning of tyrosine kinases in the early Xenopus embryo:
RT identification of Eck-related genes expressed in cranial neural crest
RT cells of the second (hyoid) arch.";
RL Dev. Dyn. 203:119-140(1995).
DR ENBL; U11730; AAA91293.1; -.
DR EMBL; U11720; AAA91283.1; -.
DR HSP; P11362; 2FGI.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004672; F:protein kinase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR ProDom; PD000001; Prot_kinase; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
FT NON_TER 1
FT NON_TER 56
SQ SEQUENCE 56 AA; 6408 MW; BD31D90C5791B45F CRC64;

Query Match 79.3%; Score 23; DB 2; Length 56;
Best Local Similarity 57.1%; Pred. No. 2.1e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VMXVAEF 7
|: ||||
Db 11 VMKIADF 17

RESULT 97
Q788X1 PRELIMINARY; PRT; 56 AA.
AC Q788X1;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE E59 protein (Fragment).
GN Name=E59;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95383727; PubMed=7655077;
RA Brandli A.W., Kirschner M.W.;
RT "Molecular cloning of tyrosine kinases in the early Xenopus embryo:
RT identification of Eck-related genes expressed in cranial neural crest
RT cells of the second (hyoid) arch.";
RL Dev. Dyn. 203:119-140(1995).
DR EMBL; U11723; AAA91286.1; -.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004672; F:protein kinase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR011009; Kinase_like.
DR ProDom; PD000001; Prot_kinase; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
FT NON_TER 1
FT NON_TER 56
SQ SEQUENCE 56 AA; 6560 MW; 09F093CAF2E32456 CRC64;

Query Match 79.3%; Score 23; DB 2; Length 56;
Best Local Similarity 57.1%; Pred. No. 2.1e+02;

```

Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
 || :||
 Db 11 VMKIADF 17

RESULT 98

Q9QW56 Q9QW56 PRELIMINARY; PRT; 57 AA.

AC Q9QW56;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Protein tyrosine kinase (Fragment).
 OS Mus sp. .
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10095;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93096484; PubMed=1281307;
 RA Glardi-Hebensreit P., Nieto M.A., Frain M., Mattei M.G.,
 R Chesler A., Wilkison D.G., Charnay P.;
 RT "An Eph-related receptor protein tyrosine kinase gene segmentally
 expressed in the developing mouse hindbrain.";
 RL Oncogene 7:2499-2506(1992).
 DR HSSP; P11362; IFGK.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0004672; F:protein kinase activity; IEA.
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
 DR InterPro; IPR011009; Kinase like.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR PROSITE; PSS0011; PROTEIN KINASE DOM; 1.
 SQ SEQUENCE 57 AA; 6632 MW; 48B1938BB3E32456 CRC64;

Query Match 79.3%; Score 23; DB 2; Length 57;
 Best Local Similarity 57.1%; Pred. No. 2.2e+02;
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
 || :||
 Db 12 VMKIADF 18

RESULT 99

Q08556 Q08556 PRELIMINARY; PRT; 57 AA.

AC Q08556;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Coturnix coturnix fibroblast growth factor receptor, clone H7
 (Fragment).
 OS Coturnix coturnix (Common Quail).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianinae;
 OC Coturnix.
 OX NCBI_TaxID=9091;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93096482; PubMed=1281306;
 RA Marcelle C., Eichmann A.;
 RT "Molecular cloning of a family of protein kinase genes expressed in
 the avian embryo.";
 RL Oncogene 7:2479-2487(1992).
 DR EMBL; X69693; CAA49363.1; -.
 DR HSSP; P11362; 2FGI.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0004672; F:protein kinase activity; IEA.
 DR GO; GO:0004872; F:receptor activity; IEA.
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.

DR InterPro; IPR011009; Kinase like.
 DR InterPro; IPR000719; Prot_kinase.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR PROSITE; PSS0011; PROTEIN_KINASE_DOM; 1.
 KW Receptor.
 FT NON_TER 1
 FT NON_TER 57 57
 SQ SEQUENCE 57 AA; 6523 MW; 8DAA6A71E3F1B45E CRC64;

Query Match 79.3%; Score 23; DB 2; Length 57;
 Best Local Similarity 57.1%; Pred. No. 2.2e+02;
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
 || :||
 Db 12 VMKIADF 18

RESULT 100

Q9P166 Q9P166 PRELIMINARY; PRT; 61 AA.

AC Q9P166;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
 DE PRO2435.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RA Zhang C., Yu Y., Zhang S., Wei H., Zhang Y., Zhou G., Bi J., Liu M.,
 RA He F.;
 RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF119881; AAF69635.1; -.
 SQ SEQUENCE 61 AA; 7467 MW; 6433950D4657C1D4 CRC64;

Query Match 79.3%; Score 23; DB 2; Length 61;
 Best Local Similarity 57.1%; Pred. No. 2.3e+02;
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
 : :|||
 Db 27 ILKVAEF 33

Search completed: June 13, 2005, 13:59:55
 Job time : 123 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2005, 13:22:36 ; Search time 118 Seconds
(without alignments)

22.943 Million cell updates/sec

Title: 09730329-59ED

Perfect score: 22

Sequence: 1 VXXAAEF 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 300 summaries

Database : A_Geneseq_16Dec04:*

1: Geneseqp1980s:*

2: Geneseqp1990s:*

3: Geneseqp2000s:*

4: Geneseqp2001s:*

5: Geneseqp2002s:*

6: Geneseqp2003as:*

7: Geneseqp2003bs:*

8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	19	86.4	146	7	ADC07972 Rice prot
2	19	86.4	243	8	ADR10461 Human pro
3	19	86.4	267	2	AAR06639 Orotidine
4	19	86.4	267	3	AAY90608 Candida t
5	19	86.4	267	5	AAU12099 Candida t
6	19	86.4	267	7	ADC44987 Yeast URA
7	19	86.4	267	7	ADC45557 Yeast URA
8	19	86.4	267	7	ADC52068 Integrati
9	19	86.4	267	7	ADF72375 Candida t
10	19	86.4	267	7	ADF11815 C. tropic
11	19	86.4	267	7	ADF11697 C. tropic
12	19	86.4	267	8	ADE64232 Integrati
13	19	86.4	267	8	ADG73846 Candida t
14	19	86.4	267	8	ADJ26614 Candida t
15	19	86.4	268	2	AAR06640 Orotidine
16	19	86.4	298	4	AAU34649 E. coli c
17	19	86.4	298	6	ABU28707 Protein e
18	19	86.4	305	4	ABG17869 Novel hum
19	19	86.4	309	2	AAW69244 Clone CK2
20	19	86.4	309	4	AAB90724 Human CK2
21	19	86.4	338	6	ABU44714 Protein e
22	19	86.4	697	4	AAB93191 Human pro
23	19	86.4	957	8	ADL90006 Gluconoba
24	19	86.4	1266	6	ABU47745 Protein e
25	19	86.4	1450	7	ADJ80131 Novel hum

1584	86.4	19	26	ABR64238	Angiogene
2214	86.4	19	27	ADN24084	Bacterial
12	81.8	18	28	ABG78424	Memapsin
12	81.8	18	29	ABG78419	Memapsin
12	81.8	18	30	ABG78418	Memapsin
13	81.8	18	31	ABG78418	Memapsin
13	81.8	18	32	ADOS2379	Human ubi
41	81.8	18	33	ADOS2379	Human ubi
44	81.8	18	34	AAB61451	Human TAN
47	81.8	18	35	AAM47175	Modular e
53	81.8	18	36	ADI21706	Novel hum
56	81.8	18	37	AAI14477	Fragment
56	81.8	18	38	ADJ12304	Peptide f
69	81.8	18	39	AAI79490	Human tes
73	81.8	18	40	ABP06455	Human ORF
83	81.8	18	41	ADB09905	Alloiooc
93	81.8	18	42	ADS21033	Bacterial
112	81.8	18	43	ABP05524	Human ORF
121	81.8	18	44	AAB29900	Human sec
137	81.8	18	45	ADF76715	Novel hum
139	81.8	18	46	AAI99815	Soybean p
144	81.8	18	47	AAU31774	Novel hum
148	81.8	18	48	ABBS5379	Lactococ
155	81.8	18	49	AAB34523	Human sec
159	81.8	18	50	ADS27899	Bacterial
162	81.8	18	51	AAU67631	Propionib
162	81.8	18	52	ABM64150	Propionib
169	81.8	18	53	AAB28213	Novel hum
179	81.8	18	54	ADI67204	Lactobaci
193	81.8	18	55	ABP29671	Streptoco
197	81.8	18	56	ABP25742	Streptoco
204	81.8	18	57	AAI34533	Porphorym
214	81.8	18	58	ABO66356	Klebsiell
221	81.8	18	59	AAI34400	Porphorym
225	81.8	18	60	AAI85820	S. pneumo
225	81.8	18	61	ABBS5926	Drosophil
228	81.8	18	62	ADJ49593	Oil-aesoc
228	81.8	18	63	ADJ49577	Oil-aesoc
231	81.8	18	64	AAW60988	Streptoco
236	81.8	18	65	AAI99822	Soybean p
236	81.8	18	66	AAB61537	Desp prot
239	81.8	18	67	AAI27923	Arabidops
240	81.8	18	68	AAI97195	E.coli ni
240	81.8	18	69	AAU29379	Novel mar
240	81.8	18	70	ABP57429	E. coli m
240	81.8	18	71	ABP57430	NADH oxid
243	81.8	18	72	AAI34534	Porphorym
247	81.8	18	73	AAU67035	Propionib
247	81.8	18	74	ABM63554	Propionib
248	81.8	18	75	AAI34401	Porphorym
249	81.8	18	76	ABBS69168	Drosophil
253	81.8	18	77	ADR01255	Farnesyl
255	81.8	18	78	ABO63928	Klebsiell
258	81.8	18	79	ABG79542	Corynebac
258	81.8	18	80	AAI92053	C glutami
266	81.8	18	81	AAI75744	Neisseria
266	81.8	18	82	AAI75743	Neisseria
266	81.8	18	83	AAI75742	Neisseria
268	81.8	18	84	AAI75953	Neisseria
268	81.8	18	85	AAU72995	Neisseria
268	81.8	18	86	ABU37163	Protein e
268	81.8	18	87	ABU38096	Protein e
269	81.8	18	88	ABP80010	N. gonorr
269	81.8	18	89	ADN61450	Human KPP
271	81.8	18	90	AAI57089	Calbindin
271	81.8	18	91	ABM78464	Human pro
271	81.8	18	92	ADB70324	Calretini
271	81.8	18	93	ADE63979	Human Pro
271	81.8	18	94	ADE63983	Human Pro
271	81.8	18	95	ADE63977	Rat Prote
271	81.8	18	96	ADE63981	Rat Prote
271	81.8	18	97	ADJ37103	Human mal
284	81.8	18	98	AAI27922	Arabidops
288	81.8	18	99	ABU44257	Protein e

99	18	81.8	290	6	ABM64812	Abm64812 Propionib	172	18	81.8	447	7	AD58093	Ad58093 Rat Prote
100	18	81.8	294	3	AAG49548	Aag49548 Arabidops	173	18	81.8	447	7	AD58099	Ad58099 Human Pro
101	18	81.8	295	7	ABO68028	AbO68028 Pseudomon	174	18	81.8	447	7	AD58107	Ad58107 Human Pro
102	18	81.8	296	3	AAG29506	Aag29506 Arabidops	175	18	81.8	447	7	AD58114	Ad58114 Rat Prote
103	18	81.8	296	4	AAM79448	Aam79448 Human pro	176	18	81.8	447	7	AD58111	Ad58111 Human Pro
104	18	81.8	299	3	AAG28024	Aag28024 Arabidops	177	18	81.8	447	7	AD58089	Ad58089 Rat Prote
105	18	81.8	312	8	ADN26970	Adn26970 Bacterial	178	18	81.8	447	7	AD585916	Ad585916 Human Pro
106	18	81.8	315	6	ABU24623	Abu24623 Protein e	179	18	81.8	447	7	AD58101	Ad58101 Rat Prote
107	18	81.8	316	3	AAG27921	Aag27921 Arabidops	180	18	81.8	447	7	AD58091	Ad58091 Human Pro
108	18	81.8	316	8	ADS24052	AdS24052 Bacterial	181	18	81.8	447	7	AD58097	Ad58097 Rat Prote
109	18	81.8	323	6	ABU41484	Abu41484 Protein e	182	18	81.8	447	7	AD58103	Ad58103 Human Pro
110	18	81.8	324	6	ABU39086	Abu39086 Protein e	183	18	81.8	447	7	AD58105	Ad58105 Rat Prote
111	18	81.8	327	8	ADR43162	Adr43162 IPT-like	184	18	81.8	447	7	AD58109	Ad58109 Rat Prote
112	18	81.8	327	8	ADR43163	Adr43163 IPT-like	185	18	81.8	447	7	AD58095	Ad58095 Human Pro
113	18	81.8	333	6	ABP57472	Abp57472 Mycobacte	186	18	81.8	447	8	ADN22192	Adn22192 Bacterial
114	18	81.8	333	6	ABU33818	Abu33818 Protein e	187	18	81.8	451	6	ABU39276	Abu39276 Protein e
115	18	81.8	333	6	ABU36897	Abu36897 Protein e	188	18	81.8	453	3	AAV68799	Aay68799 Amino aci
116	18	81.8	333	6	ABU34521	Abu34521 Protein e	189	18	81.8	467	8	ADG63848	Adg63848 Aspergill
117	18	81.8	333	6	ABU35991	Abu35991 Protein e	190	18	81.8	475	2	AAW35724	Aaw35724 T. lanugi
118	18	81.8	335	4	AAU37753	Aau37753 Streptoco	191	18	81.8	475	2	AAW27384	Aaw27384 Thermomyc
119	18	81.8	335	6	ABU01294	Abu01294 S. pneumo	192	18	81.8	475	2	AAV39898	Aay39898 T. lanugi
120	18	81.8	335	6	ABP81576	Abp81576 Streptoco	193	18	81.8	475	2	AAW94306	Aaw94306 Thermomyc
121	18	81.8	335	6	ABU45990	Abu45990 Protein e	194	18	81.8	475	4	ABN83952	Aab83952 Amino aci
122	18	81.8	335	8	ADK48584	Adk48584 Streptoco	195	18	81.8	481	2	AAW44220	Aar44220 Threonine
123	18	81.8	338	2	AAW60797	Aar60797 Brevibact	196	18	81.8	481	4	AAW79721	Aab79721 Corynebac
124	18	81.8	338	4	AAW79944	Aab79944 Corynebac	197	18	81.8	481	4	AAW93232	Aag93232 C glutami
125	18	81.8	338	4	AAW79776	Aab79776 Corynebac	198	18	81.8	482	4	AAU31776	Aau31776 Novel hum
126	18	81.8	338	4	AAW79153	Aab79153 Corynebac	199	18	81.8	482	2	AAV33671	Aay33671 B. baessia
127	18	81.8	338	4	AAW931205	Aag931205 C glutami	200	18	81.8	495	1	AAW81325	Aap81325 Threonine
128	18	81.8	343	8	ADG27879	Adg27879 Bacterial	201	18	81.8	495	7	ADN13447	Adn13447 C. glutam
129	18	81.8	344	4	ABW65507	Abw65507 Drosophil	202	18	81.8	495	8	ADN22547	Adn22547 Bacterial
130	18	81.8	345	4	ABG26135	Abg26135 Novel hum	203	18	81.8	495	8	ADN22548	Adn22548 Bacterial
131	18	81.8	345	8	ADG22787	Adg22787 Bacterial	204	18	81.8	495	8	ADS23129	AdS23129 Bacterial
132	18	81.8	347	3	AAG18096	Aag18096 Arabidops	205	18	81.8	499	4	ABG62187	Abg62187 Drosophil
133	18	81.8	347	4	AAG98876	Aag98876 E. coli g	206	18	81.8	500	3	AGI18095	Aag18095 Arabidops
134	18	81.8	349	8	ADJ50144	Adj50144 Oil-assoc	207	18	81.8	501	6	ABU25758	Abu25758 Protein e
135	18	81.8	359	4	ABG61975	Abg61975 S. avermi	208	18	81.8	513	2	AAW32904	Aar32904 Human BMP
136	18	81.8	363	5	ABW47393	Abw47393 Listeria	209	18	81.8	513	2	AAW07311	Aar07311 Human BMP
137	18	81.8	366	6	ABU27971	Abu27971 Protein e	210	18	81.8	513	2	AAW36735	Aar36735 Human BMP
138	18	81.8	368	8	ADS28440	AdS28440 Bacterial	211	18	81.8	513	2	AAW47264	Aar47264 Pre-pro B
139	18	81.8	371	4	AAG92937	Aag92937 C glutami	212	18	81.8	513	3	ABW43307	Aab43307 Human ORF
140	18	81.8	377	3	AGG36174	Aag36174 Arabidops	213	18	81.8	513	4	AAW97702	Aau97702 Human tra
141	18	81.8	380	3	AGG36173	Aag36173 Arabidops	214	18	81.8	513	4	AAU14473	Aau14473 Human nov
142	18	81.8	380	4	ABG06781	Abg06781 Novel hum	215	18	81.8	513	7	ADB46399	AdB46399 Human Pro
143	18	81.8	380	7	ADE09368	AdE09368 Novel pro	216	18	81.8	513	7	ADG63743	AdG63743 Human Pro
144	18	81.8	387	3	AAV79489	Aay79489 Human tes	217	18	81.8	513	7	ADN95841	Adn95841 Human BEC
145	18	81.8	391	8	ADN04908	Adn04908 Antipgori	218	18	81.8	513	8	ADG14171	AdG14171 Human ful
146	18	81.8	391	8	ADO20123	Ado20123 Human PRO	219	18	81.8	513	8	ADH11583	Adh11583 Human bon
147	18	81.8	391	8	ADM32277	Adm32277 S. pneumo	220	18	81.8	513	8	ADH80791	Adh80791 Human pol
148	18	81.8	396	8	ADS29673	AdS29673 Bacterial	221	18	81.8	513	8	ADK90641	Adk90641 Human BMP
149	18	81.8	397	3	AAW81772	Aay81772 Streptoco	222	18	81.8	525	5	ADR41523	Adr41523 Human CD-
150	18	81.8	397	6	ABU02769	Abu02769 S. pneumo	223	18	81.8	526	6	ABR53376	AbR53376 Protein s
151	18	81.8	397	6	ABU46324	Abu46324 Protein e	224	18	81.8	526	7	ADK61992	Adk61992 Disease t
152	18	81.8	397	8	ADK47887	Adk47887 Streptoco	225	18	81.8	526	8	ADK43482	AdK43482 Bacterial
153	18	81.8	400	8	ADR95965	Adr95965 Novel S.	226	18	81.8	532	5	ABW77577	Abw77577 Physcomit
154	18	81.8	402	6	ABU44585	Abu44585 Protein e	227	18	81.8	532	6	ABG74202	Abg74202 P. patens
155	18	81.8	403	7	ABO63525	AbO63525 Klebsiell	228	18	81.8	560	6	ADA55710	Ada55710 Human pro
156	18	81.8	405	8	ADJ49370	Adj49370 Oil-assoc	229	18	81.8	567	5	ADR41405	Adr41405 Human CD-
157	18	81.8	412	8	ADS23674	AdS23674 Bacterial	230	18	81.8	585	5	ADJ49879	Adj49879 Oil-assoc
158	18	81.8	418	8	ADN24951	Adn24951 Bacterial	231	18	81.8	592	8	ADJ49381	Adj49381 Oil-assoc
159	18	81.8	420	3	AAG28023	Aag28023 Arabidops	232	18	81.8	602	5	ABW49009	Abw49009 Listeria
160	18	81.8	434	2	AAW01520	Aaw01520 Mega prot	233	18	81.8	605	3	ABW36443	Abw36443 Strawberry
161	18	81.8	434	2	AAW27454	Aaw27454 Magnetoosp	234	18	81.8	605	3	AAW79660	Aay79660 Strawberry
162	18	81.8	440	3	AAW69564	Aay69564 Thermomyc	235	18	81.8	605	8	ADJ49381	Adj49381 Oil-assoc
163	18	81.8	440	3	ABW20522	AbW20522 Thermomyc	236	18	81.8	605	8	ADJ48993	Adj48993 Oil-assoc
164	18	81.8	440	8	ADL73088	AdL73088 Thermomyc	237	18	81.8	607	3	AAV67411	Aay67411 Arabidops
165	18	81.8	443	6	ABA42806	AbA42806 Human ORF	238	18	81.8	607	3	AGC36172	Aag36172 Arabidops
166	18	81.8	443	6	ADA36229	Ada36229 Acinetoba	239	18	81.8	607	5	ABW93818	Abw93818 Herbicida
167	18	81.8	443	7	ADE55920	AdE55920 Human Pro	240	18	81.8	607	8	ADJ48549	Adj48549 Oil-assoc
168	18	81.8	443	7	ADE55922	AdE55922 Rat Prote	241	18	81.8	607	8	ADJ50436	Adj50436 Oil-assoc
169	18	81.8	443	7	ADE55924	AdE55924 Human Pro	242	18	81.8	607	8	ADJ49327	Adj49327 Oil-assoc
170	18	81.8	443	7	ADE55918	AdE55918 Rat Prote	243	18	81.8	608	8	ADJ48509	Adj48509 Oil-assoc
171	18	81.8	447	7	ADD18998	Add18998 Human dis	244	18	81.8	616	6	ABJ19410	Abj19410 Human int

245 18 81.8 626 6 ABU41118 Protein e
246 18 81.8 627 7 ADB651129 Human pro
247 18 81.8 628 8 ADB651129 Human pro
248 18 81.8 629 9 ADB651129 Human pro
249 18 81.8 630 10 ADB651129 Human pro
250 18 81.8 631 11 ADB651129 Human pro
251 18 81.8 632 12 ADB651129 Human pro
252 18 81.8 633 13 ADB651129 Human pro
253 18 81.8 634 14 ADB651129 Human pro
254 18 81.8 635 15 ADB651129 Human pro
255 18 81.8 636 16 ADB651129 Human pro
256 18 81.8 637 17 ADB651129 Human pro
257 18 81.8 638 18 ADB651129 Human pro
258 18 81.8 639 19 ADB651129 Human pro
259 18 81.8 640 20 ADB651129 Human pro
260 18 81.8 641 21 ADB651129 Human pro
261 18 81.8 642 22 ADB651129 Human pro
262 18 81.8 643 23 ADB651129 Human pro
263 18 81.8 644 24 ADB651129 Human pro
264 18 81.8 645 25 ADB651129 Human pro
265 18 81.8 646 26 ADB651129 Human pro
266 18 81.8 647 27 ADB651129 Human pro
267 18 81.8 648 28 ADB651129 Human pro
268 18 81.8 649 29 ADB651129 Human pro
269 18 81.8 650 30 ADB651129 Human pro
270 18 81.8 651 31 ADB651129 Human pro
271 18 81.8 652 32 ADB651129 Human pro
272 18 81.8 653 33 ADB651129 Human pro
273 18 81.8 654 34 ADB651129 Human pro
274 18 81.8 655 35 ADB651129 Human pro
275 18 81.8 656 36 ADB651129 Human pro
276 18 81.8 657 37 ADB651129 Human pro
277 18 81.8 658 38 ADB651129 Human pro
278 18 81.8 659 39 ADB651129 Human pro
279 18 81.8 660 40 ADB651129 Human pro
280 18 81.8 661 41 ADB651129 Human pro
281 18 81.8 662 42 ADB651129 Human pro
282 18 81.8 663 43 ADB651129 Human pro
283 18 81.8 664 44 ADB651129 Human pro
284 18 81.8 665 45 ADB651129 Human pro
285 18 81.8 666 46 ADB651129 Human pro
286 18 81.8 667 47 ADB651129 Human pro
287 18 81.8 668 48 ADB651129 Human pro
288 18 81.8 669 49 ADB651129 Human pro
289 18 81.8 670 50 ADB651129 Human pro
290 18 81.8 671 51 ADB651129 Human pro
291 18 81.8 672 52 ADB651129 Human pro
292 18 81.8 673 53 ADB651129 Human pro
293 18 81.8 674 54 ADB651129 Human pro
294 18 81.8 675 55 ADB651129 Human pro
295 18 81.8 676 56 ADB651129 Human pro
296 18 81.8 677 57 ADB651129 Human pro
297 18 81.8 678 58 ADB651129 Human pro
298 18 81.8 679 59 ADB651129 Human pro
299 18 81.8 680 60 ADB651129 Human pro
300 18 81.8 681 61 ADB651129 Human pro

ALIGNMENTS

RESULT 1
ADC07972
ID ADC07972 standard; protein; 146 AA.
XX
AC ADC07972;
XX
DT 18-DEC-2003 (first entry)
XX
DE Rice protein sequence Seq ID238 related to grain filling.
XX
KW plant biotechnology; carbohydrate synthesis; carbohydrate metabolism;
KW carbohydrate degradation; carbohydrate; plant grain; grain filling; corn;
KW

tomato; banana; canola; cotton; peanut; sorghum; tobacco; sugarbeet;
wheat; rice; protein; oil; starch; fibre; moisture content; cereal grain;
gene; da; plant.
Oryza sativa.
WO2003000905-A2.
03-JAN-2003.
21-JUN-2002; 2002WO-IB002450.
22-JUN-2001; 2001US-0300112P.
26-SEP-2001; 2001US-0325277P.
20-DEC-2001; 2001US-0342327P.
(SYGN) SYNGENTA PARTICIPATIONS AG.
Zhu T, Cheng W, Briggs S, Cooper B, Goff SA, Moughamer T;
Glazebrook J, Katagiri F, Kreps J, Provart N, Ricke D;
WPI: 2003-229341/22.
N-PSDB; ADC07971.
New plant genes encoding polypeptides having an activity involved in or
associated with the synthesis, metabolism or degradation of carbohydrates
in the plant grain useful in generating plants having improved
nutritional properties.
Claim 15; SEQ ID NO 238; 130pp; English.
This invention, in the area of plant biotechnology, relates to novel
polynucleotides comprising a nucleotide sequence encoding a protein which
is involved in or associated with the synthesis, metabolism or
degradation of carbohydrates in the plant grain and the expression of
which is up-regulated during grain filling. The plant is selected from
corn, tomato, banana, canola, cotton, peanut, sorghum, tobacco,
sugarbeet, wheat, and rice. The invention may be useful for the
improvement of protein, oil, starch, fibre and moisture content of the
cereal grains. In addition, carbohydrate levels may be modified to a more
desirable level using the present invention. The present sequence is the
amino acid sequence of a rice protein of the invention. Note: The
sequence data for this patent did not form part of the printed
specification, but was obtained in electronic format directly from WIPO
at ftp.wipo.int/pub/publishedpct_sequences.
Query Match 86.4%; Score 19; DB 7; Length 146;
Best Local Similarity 57.1%; Pred. No. 9.1e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 VXXAXAF 7
DB 2 VAATAEF 8
Sequence 146 AA;
RESULT 2
ADRI0461
ID ADRI0461 standard; protein; 243 AA.
XX
AC ADRI0461;
XX
DT 04-NOV-2004 (first entry)
XX
DE Human protein useful for treating neurological disease Seq 3967.
XX
KW human; oligo-capping method; diagnostic marker; gene therapy;
KW osteoporosis; neurological disease; Alzheimer's disease;
KW Parkinson's disease; dementia; short memory; cancer;
KW sense or motor function; emotional reaction; fear response; panic;
KW osteopathic; neuroprotective; nootropic; antiparkinsonian; cytostatic;
KW tranquilliser.

XX OS Homo sapiens.
 XX PF EP1447413-A2.
 XX XX 18-AUG-2004.
 XX PF 12-FEB-2004; 2004EP-00003145.
 XX PR 14-FEB-2003; 2003JP-00102207.
 XX PR 09-MAY-2003; 2003JP-00131452.
 XX PA (REAS-) RES ASSOC BIOTECHNOLOGY.
 XX XX Isogai T, Yamamoto J, Nishikawa T, Isono Y, Sugiyama T, Otsuki T;
 PI Wakamatsu A, Ishii S, Nagai K, Irie R;
 XX WPI: 2004-583265/57.
 DR N-PSDB; ADR10422.
 XX New 1995 cDNA, useful for treating osteoporosis, neurological diseases,
 PT Alzheimer's diseases, Parkinson's diseases, dementia and various cancers.
 XX Claim 1; SEQ ID NO 3967; 2686pp; English.
 XX This invention relates to novel, isolated full length human cDNA
 CC molecules and the encoded proteins thereof. Specifically, it refers to
 CC cDNA clones obtained by an oligo-capping method, where none of these
 CC clones are identical to any known human mRNAs. The present invention
 CC describes an immunoassay to identify agonists and antagonists, as well as
 CC antibodies, antisense molecules and siRNAs that can all be used to bind
 CC to and modulate expression of the cDNA molecules. As such, these
 CC molecules are useful for diagnostic markers or therapeutic targets for
 CC the various diseases or morbid states. In particular, they are useful in
 CC gene therapy for treating osteoporosis, neurological disease, Alzheimer's
 CC disease, Parkinson's disease, dementia, short memory and various cancers,
 CC as well as for maintaining equilibrium of sense or motor function, and
 CC for treating emotional reaction, fear response and panic. Accordingly,
 CC they exhibit osteopathic, neuroprotective, neurotropic, antiparkinsonian,
 CC cytotatic and tranquiliser activities. This polypeptide is a protein
 CC encoded by a full length human cDNA sequence of the invention. NOTE: This
 CC sequence is not given in the sequence listing of the specification but
 CC can be obtained on CD-ROM from the European Patent Office, Vienna Sub-
 CC office.
 XX XX Sequence 243 AA;
 SQ Query Match 86.4%; Score 19; DB 8; Length 243;
 Best Local Similarity 57.1%; Pred. No. 1.6e+03;
 Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 VXXXXAEF 7
 DB 201 VATAAEF 207
 RESULT 3
 ID AAR06639
 XX AAR06639 standard; protein; 267 AA.
 XX AC AAR06639;
 XX DT 09-JAN-1991 (first entry)
 XX DE Orotidine-5'-phosphate decarboxylase encoded by the URA3 A gene.
 XX KW Uracil biosynthetic pathway; orotidine-5'-phosphate decarboxylase;
 XX URA3 A gene; auxotrophic; mutagenesis.
 XX OS Candida tropicalis.
 XX PN WO9009449-A.
 XX

PD XX 23-AUG-1990.
 XX PF 08-FEB-1989; 89US-00308481.
 XX PR 08-FEB-1989; 89US-00308481.
 XX PR 27-JUL-1989; 89US-00386837.
 XX PA (HENK) HENKEL RES CORP.
 XX Cregg JM, Gleeson MA, Picataggio S;
 PI WPI: 1990-275138/36.
 DR N-PSDB; AAQ05864.
 XX Candida tropicalis transformation system - using host auxotrophic in an
 PT amino acid, purine or pyrimidine pathway as result of mutagenesis.
 XX Disclosure; Fig 3; 43pp; English.
 XX The gene product (and opt. the URA3 B gene product - AAR06640) is used
 CC to complement an auxotrophic host to prototrophy. The host may be
 CC defective in orotidine-5'-phosphate decarboxylase activity. See also
 CC AAQ05864-Q05866
 XX XX Sequence 267 AA;
 SQ Query Match 86.4%; Score 19; DB 2; Length 267;
 Best Local Similarity 57.1%; Pred. No. 1.8e+03;
 Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 VXXXXAEF 7
 DB 39 VTTTAEF 45
 RESULT 4
 ID AAY90608
 XX AAY90608 standard; protein; 267 AA.
 XX AC AAY90608;
 XX DT 12-SEP-2003 (revised)
 DT 21-AUG-2000 (first entry)
 XX Candida tropicalis URA3A protein.
 DE URA3; URA integration vector; cytochrome P450; NADPH reductase;
 KW monooxygenase; CYP52A; CPR.
 XX OS Candida tropicalis; ATCC20366.
 XX WO200020566-A2.
 PN 13-APR-2000.
 XX 10-SEP-1999; 99WO-US020797.
 XX 05-OCT-1998; 98US-0103099P.
 PR 10-MAR-1999; 99US-0123555P.
 XX (HENK) HENKEL CORP.
 XX Wilson CR, Craft DL, Eirich LD, Eshoo M, Madduri KM, Cornett CA;
 PI Brenner AA, Tang M, Loper JC, Gleeson M;
 XX WPI: 2000-317711/27.
 DR N-PSDB; AAA30578.
 XX Cytochrome P450 nicotine adenine dinucleotide phosphate oxidoreductase
 PT and cytochrome P450 monooxygenase nucleic acids and encoded proteins,
 PT useful for overproducing dicarboxylic acids.
 XX Example 15; Fig 23; 200pp; English.
 PS

XX The invention relates to 12 novel genomic DNA sequences and proteins
CC which are components of the omega hydroxylase complex of Candida
CC tropicalis ATCC 20366. The DNA sequences (AA30566-A30577) respectively
CC encode cytochrome P450 NADPH oxidoreductases CPRA and CPRB (AA190596,
CC AA190597) and cytochrome P450 monooxygenases CYP52A1A, CYP52A2A,
CC CYP52A2B, CYP52A3A, CYP52A3B, CYP52A5A, CYP52A5B, CYP52A8A, CYP52A8B and
CC CYP52D4A (AA190598-Y90607). Of the cytochrome P450 DNAs isolated, six are
CC unique CYP genes and four are potential alleles. The omega hydroxylase
CC complex is a membrane-bound enzyme complex found in certain yeasts which
CC catalyses the first step in the omega-oxidation of fatty acids or
CC alkanes, this being primary oxidation of the terminal methyl group. Such
CC yeasts, which include members of the genus *Candida*, excrete alpha-omega-
CC dicarboxylic acids when alkanes or fatty acids are used as the carbon
CC source. The products of the P450 genes CYP52A1, CYP52A2 and CYP52A5 were
CC identified as playing a greater role in the omega-oxidation of long chain
CC fatty acids via a novel quantitative competitive reverse transcription-
CC PCR (QC-RT-PCR). This assay quantifies the amount of target mRNA in a
CC sample and may be used for discriminating members of a gene family, such
CC as the CYP gene family. Organisms containing the target gene are cultured
CC on an organic substrate which causes upregulation of that gene. The total
CC RNA is then extracted and mixed with a known amount of competitor RNA,
CC which is similar to the target mRNA but has fewer nucleotides. RT-PCR
CC reactions are performed using increasing amounts of competitor RNA and
CC the point at which the amount of the corresponding target DNA is equal to
CC the amount of the corresponding competitor DNA is determined. The CPR and
CC CYP nucleic acids may be transformed into a suitable host so that the
CC host overexpresses the corresponding proteins. Such host cells will
CC overproduce dicarboxylic acids. The dicarboxylic acids thus produced find
CC application as thermoplastics, plasticising agents, lubricants, hydraulic
CC fluids, agricultural chemicals, pharmaceuticals, dyes, surfactants,
CC adhesives and fragrances. The CPR and CYP nucleic acids and proteins
CC enable inexpensive large scale production of industrially useful
CC dicarboxylic acids. The present sequence represents the *Candida*
CC tropicalis ATCC20366 URA3 protein. The gene encoding it was used in an
CC exemplification of the invention to construct a URA integration vector.
CC (Updated on 12-SEP-2003 to standardise OS field)

XX
SQ Sequence 267 AA;

Query Match 86.4%; Score 19; DB 3; Length 267;
Best Local Similarity 57.1%; Pred. No. 1.8e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXXAEEF 7
Db 39 VTTTAEF 45

RESULT 5
AAU12099
ID AAU12099 standard; protein; 267 AA.
XX
AC AAU12099;
XX
DT 29-AUG-2003 (revised)
DT 23-APR-2002 (first entry)
XX
DE *Candida tropicalis* URA3A protein.
XX
KW CPRA; CYP52A1A; CYP52A2A; CYP52A2B; CYP52A3A; CYP52A3B; CYP52A5A;
KW CYP52A5B; CYP52A8A; CYP52A8B; CYP52D4A; URA3A; cytochrome P450;
KW NADPH2 reductase; omega-hydroxylase complex; dicarboxylic acid; enzyme.
XX
OS *Candida tropicalis*; 20336.
XX
PN US6331420-B1.
XX
PD 18-DEC-2001.
XX
PF 30-APR-1999; 99US-00302620.
XX
XX 01-MAY-1998; 98US-0083798P.

PR 05-OCT-1998; 98US-0103099P.
PR 10-MAR-1999; 99US-0123555P.
XX
PA (WILS/) WILSON C R.
PA (CRAF/) CRAFT D L.
PA (EIRI/) EIRICH L D.
PA (ESHO/) ESHOO M.
PA (MADD/) MADDURI K M.
PA (CORN/) CORNETT C A.
PA (BREN/) BRENNER A A.
PA (TANG/) TANG M.
PA (LOPE/) LOPEZ J C.
PA (GLEE/) GLEESON M.
XX
PI Wilson CR, Craft DL, Eirich LD, Eshoo M, Madduri KM, Cornett CA;
PI Brenner AA, Tang M, Loper JC, Gleeson M;
XX
DR WPI; 2002-138983/18.
DR N-PSDB; ABK31897.
XX
PT Novel isolated nucleic acid encoding cytochrome P450 and NADPH reductase
PT enzymes of omega-hydroxylase complex of *Candida tropicalis*, useful for
PT increasing production of dicarboxylic acids.
XX
PS Example 15; Fig 23; 173pp; English.
XX
CC The present invention relates to the isolation of *Candida tropicalis*
CC 20336 novel genes (CPRA, CPRB, CYP52A1A, CYP52A2A, CYP52A2B, CYP52A3A,
CC CYP52A3B, CYP52A5A, CYP52A5B, CYP52A8A, CYP52A8B, CYP52D4A and URA3A)
CC which encode cytochrome P450 and NADPH2 reductase enzymes of the omega-
CC hydroxylase complex. Also disclosed are vectors containing these genes
CC and methods of producing these enzymes. The genes and vectors are useful
CC for increasing production of a dicarboxylic acid by providing a host cell
CC having a naturally occurring number of the genes of the invention and
CC increasing in the host cell, the number of genes encoding these enzymes.
CC The present sequence represents an enzyme of the present invention.
CC (Updated on 29-AUG-2003 to standardise OS field)

XX
SQ Sequence 267 AA;

Query Match 86.4%; Score 19; DB 5; Length 267;
Best Local Similarity 57.1%; Pred. No. 1.8e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXXAEEF 7
Db 39 VTTTAEF 45

RESULT 6
ADC44987
ID ADC44987 standard; protein; 267 AA.
XX
AC ADC44987;
XX
DT 18-DEC-2003 (first entry)
XX
DE Yeast URA3A protein.
XX
KW Gene; enzyme; yeast; omega oxygenase complex;
KW cytochrome P450 monooxygenase; CYP; NADPH reductase enzymes; CPR; CPRA;
KW CPRB; CYP52A1A; CYP52A2A; CYP52A2B; CYP52A3A; CYP52A3B; CYP52A5A;
KW CYP52A5B; CYP52A8A; CYP52A8B; CYP52D4A; dicarboxylic acid; diester;
KW polymer; thermoplastic; plasticising agent; lubricant; hydraulic fluid;
KW agricultural chemical; pharmaceutical; dye; surfactant; adhesive; URA3.
XX
OS *Candida tropicalis*.
XX
PN US2003049821-A1.
XX
PD 13-MAR-2003.
XX
PF 03-MAY-2002; 2002US-00138838.

XX 01-MAY-1998; 98US-0083798P.
 PR 05-OCT-1998; 98US-0103099P.
 PR 10-MAR-1999; 99US-0123555P.
 PR 30-APR-1999; 99US-00302620.
 PR 12-OCT-2001; 2001US-00976800.
 XX (WILS/) WILSON C R.
 PA (CRAF/) CRAFT D L.
 PA (EIRI/) EIRICH L D.
 PA (ESHO/) ESHOO M.
 PA (MADD/) MADDURI K M.
 PA (CORN/) CORNETT C A.
 PA (BREN/) BRENNER A A.
 PA (TANG/) TANG M.
 PA (LOPE/) LOPEZ J C.
 PA (GLEE/) GLEESON M.
 XX Wilson CR, Craft DL, Eirich LD, Eshoo M, Madduri KM, Cornett CA;
 PI Brenner AA, Tang M, Loper JC, Gleeson M;
 XX WPI; 2003-777150/73.
 DR N-PSDB; ADC44986.
 XX New nucleic acid encoding cytochrome P450 and NADPH reductase enzymes
 PT (e.g. CPRA, CPRB or CYP52A1A), useful for producing dicarboxylic acids
 PT that may be utilized as industrial intermediates in manufacturing
 PT diesters and polymers.
 XX Example 15; SEQ ID NO 106; 196pp; English.
 XX The invention relates to an isolated nucleic acid selected encoding
 CC Candida tropicalis omega oxygenase complex enzymes (cytochrome P450
 CC monooxygenase (CYP) and NADPH reductase enzymes (CPR) designated CPRA,
 CC CPRB, CYP52A1A, CYP52A2A, CYP52A2B, CYP52A3A, CYP52A3B, CYP52A5A,
 CC CYP52A5B, CYP52A8A, CYP52A8B or CYP52D4A) or their coding regions. Also
 CC included are the CPR/CYP proteins, a vector comprising the nucleic acid
 CC cited above, a host cell transfected or transformed with the above
 CC nucleic acid, producing the amount of target mRNA in a sample, increasing
 CC family by quantifying the amount of target mRNA in a sample, increasing
 CC production of a dicarboxylic acid and increasing the production of the
 CC proteins cited above. The host cell is C. tropicalis is specifically
 CC H5343 ura-. The nucleic acid is useful for producing dicarboxylic acids
 CC that may be utilized as industrial intermediates in the manufacture of
 CC diesters and polymers (e.g. as thermoplastics, plasticising agents,
 CC lubricants, hydraulic fluids, agricultural chemicals, pharmaceuticals,
 CC dyes, surfactants or adhesives). The present sequence is encoded by the
 CC yeast URA3 gene which is targeted for disruption by an integrative vector
 CC carrying a CYP or CPR gene.
 XX Sequence 267 AA;
 SQ
 Query Match 86.4%; Score 19; DB 7; Length 267;
 Best Local Similarity 57.1%; Pred. No. 1.8e+03;
 Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 VXXXAEF 7
 Db 39 VTTTAEF 45
 RESULT 7
 ADC4557
 ID ADC45557 standard; protein; 267 AA.
 XX ADC45557;
 AC
 DT 18-DEC-2003 (first entry)
 XX Yeast URA3A protein.
 DE
 XX Gene; enzyme; yeast; omega oxygenase complex;
 KW cytochrome P450 monooxygenase; CYP; NADPH reductase enzymes; CPR; CPRA;
 KW

KW CPRB; CYP52A1A; CYP52A2A; CYP52A2B; CYP52A3A; CYP52A3B; CYP52A5A;
 KW CYP52A5B; CYP52A8A; CYP52A8B; CYP52D4A; dicarboxylic acid; diester;
 KW polymer; thermoplastic; plasticising agent; lubricant; hydraulic fluid;
 KW agricultural chemical; pharmaceutical; dye; surfactant; adhesive; URA3.
 XX Candida tropicalis.
 XX US2003049822-A1.
 PN 13-MAR-2003.
 XX 03-MAY-2002; 2002US-00139031.
 XX 01-MAY-1998; 98US-0083798P.
 PR 05-OCT-1998; 98US-0103099P.
 PR 10-MAR-1999; 99US-0123555P.
 PR 30-APR-1999; 99US-00302620.
 PR 12-OCT-2001; 2001US-00976800.
 XX (WILS/) WILSON C R.
 PA (CRAF/) CRAFT D L.
 PA (EIRI/) EIRICH L D.
 PA (ESHO/) ESHOO M.
 PA (MADD/) MADDURI K M.
 PA (CORN/) CORNETT C A.
 PA (BREN/) BRENNER A A.
 PA (TANG/) TANG M.
 PA (LOPE/) LOPEZ J C.
 PA (GLEE/) GLEESON M.
 XX Wilson CR, Craft DL, Eirich LD, Eshoo M, Madduri KM, Cornett CA;
 PI Brenner AA, Tang M, Loper JC, Gleeson M;
 XX WPI; 2003-765370/72.
 DR N-PSDB; ADC45556.
 XX New nucleic acid encoding cytochrome P450 and NADPH reductase enzymes
 PT (e.g. CPRA, CPRB or CYP52A1A), useful for producing dicarboxylic acids
 PT that may be utilized as industrial intermediates in manufacturing
 PT diesters and polymers.
 XX Example 15; SEQ ID NO 106; 196pp; English.
 XX The invention relates to an isolated nucleic acid selected encoding
 CC Candida tropicalis omega oxygenase complex enzymes (cytochrome P450
 CC monooxygenase (CYP) and NADPH reductase enzymes (CPR) designated CPRA,
 CC CPRB, CYP52A1A, CYP52A2A, CYP52A2B, CYP52A3A, CYP52A3B, CYP52A5A,
 CC CYP52A5B, CYP52A8A, CYP52A8B or CYP52D4A) or their coding regions. Also
 CC included are the CPR/CYP proteins, a vector comprising the nucleic acid
 CC cited above, a host cell transfected or transformed with the above
 CC nucleic acid, producing the amount of target mRNA in a sample, increasing
 CC family by quantifying the amount of target mRNA in a sample, increasing
 CC production of a dicarboxylic acid and increasing the production of the
 CC proteins cited above. The host cell is C. tropicalis is specifically
 CC H5343 ura-. The nucleic acid is useful for producing dicarboxylic acids
 CC that may be utilized as industrial intermediates in the manufacture of
 CC diesters and polymers (e.g. as thermoplastics, plasticising agents,
 CC lubricants, hydraulic fluids, agricultural chemicals, pharmaceuticals,
 CC dyes, surfactants or adhesives). The present sequence is encoded by the
 CC yeast URA3 gene which is targeted for disruption by an integrative vector
 CC carrying a CYP or CPR gene.
 XX Sequence 267 AA;
 SQ
 Query Match 86.4%; Score 19; DB 7; Length 267;
 Best Local Similarity 57.1%; Pred. No. 1.8e+03;
 Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 VXXXAEF 7
 Db 39 VTTTAEF 45

RESULT 8
ADE52068
ID ADE52068 standard; protein; 267 AA.
XX
AC ADE52068;
XX
XX 29-JAN-2004 (first entry)
XX
DE Integration vector pURA2in URA3A gene encoded protein.
XX
XX Yeast; enzyme; cytochrome P450; CYP; NADPH reductase; CPR;
KW omega-hydroxylase complex; omega-oxidation; long chain fatty acid; URA3.
XX
XX Candida tropicalis.
XX
XX US2003073220-A1.
XX
PD 17-APR-2003.
XX
PF 03-MAY-2002; 2002US-00138916.
XX
PR 01-MAY-1998; 98US-0083798P.
PR 05-OCT-1998; 98US-0103099P.
PR 10-MAR-1999; 99US-0123555P.
PR 30-APR-1999; 99US-00302620.
PR 12-OCT-2001; 2001US-00976800.
XX
XX (WILS/) WILSON C R.
PA (CRAF/) CRAFT D L.
PA (EIRI/) EIRICH L D.
PA (ESHO/) ESHOO M.
PA (MADD/) MADDURI K M.
PA (CORN/) CORNETT C A.
PA (BREN/) BRENNER A A.
PA (TANG/) TANG M.
PA (LOPE/) LOPER J C.
PA (GLEE/) GLEESON M.
XX
XX Wilson CR, Craft DL, Eirich LD, Eshoo M, Madduri KM, Cornett CA;
PI Brenner AA, Tang M, Loper JC, Gleeson M;
PI
DR WPI; 2003-625522/59.
DR N-PSDB; ADE52067.
XX
XX
XX New cytochrome P450 and NADPH oxidoreductase, i.e. CPR and CYP, genes and
PT proteins, useful for discriminating members of a gene family by
PT quantifying the amount of target mRNA in a sample, or for omega-oxidation
PT of long chain fatty acids.
XX
XX Example 15; SEQ ID NO 106; 194pp; English.
XX
XX The invention relates to isolated nucleic acids encoding cytochrome P450
CC (CYP) and NADPH reductase (CPR) enzymes of the omega-hydroxylase complex
CC of Candida tropicalis. Also included are the CYP and CPR proteins
CC comprising CPRA, CPRB, CYP52A1A, CYP52A2A, CYP52A2B, CYP52A3A, CYP52A3B,
CC CYP52A5A, CYP52A5B, CYP52A8A, CYP52A8B, or CYP52D4A), a vector comprising
CC any one of the nucleic acid sequences cited above, a host cell
CC transfected or transformed with the nucleic acid, methods of producing
CC the CPR or CYP proteins, a method for discriminating members of a gene
CC family by quantifying the amount of target mRNA in a sample and methods
CC for increasing the production of a dicarboxylic acid, (or the CPR/CYP
CC proteins). The CPR and CYP genes and proteins are useful for
CC discriminating members of a gene family by quantifying the amount of
CC target mRNA in a sample, for increasing production of a dicarboxylic
CC acid, or for omega-oxidation of long chain fatty acids. The present
CC sequence is encoded by the URA3 gene, used in an integration vector for
CC CYP or CPR genes.
XX
XX Sequence 267 AA;
SQ

Query Match 86.4%; Score 19; DB 7; Length 267;
Best Local Similarity 57.1%; Pred. No. 1.8e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXXAEEF 7
Db 39 VTTTAEF 45
RESULT 9
ADF72375
ID ADF72375 standard; protein; 267 AA.
XX
AC ADF72375;
XX
DT 12-FEB-2004 (first entry)
XX
XX Candida tropicalis associated polypeptide seq id 106.
DE
XX CYP52A2B; cytochrome P450; NADH reductase; dicarboxylic acid production;
KW organic substrate oxidation; fatty acid oxidation;
KW gene integration vector; CYP; CPR.
XX
XX Candida tropicalis.
XX
PN US2003077795-A1.
XX
PD 24-APR-2003.
XX
PF 12-OCT-2001; 2001US-00976800.
XX
PR 10-MAR-1999; 99US-0123555P.
XX
XX (WILS/) WILSON C R.
PA (CRAF/) CRAFT D L.
PA (EIRI/) EIRICH L D.
PA (ESHO/) ESHOO M.
PA (MADD/) MADDURI K M.
PA (CORN/) CORNETT C A.
PA (BREN/) BRENNER A A.
PA (TANG/) TANG M.
PA (LOPE/) LOPER J C.
PA (GLEE/) GLEESON M.
XX
XX Wilson CR, Craft DL, Eirich LD, Eshoo M, Madduri KM, Cornett CA;
PI Brenner AA, Tang M, Loper JC, Gleeson M;
PI
XX WPI; 2003-810780/76.
DR
XX New nucleic acids encoding a CYP52A2B protein useful for increasing the
PT production of dicarboxylic acid for oxidizing organic substrates such as
PT fatty acids.
XX
XX Disclosure; SEQ ID NO 106; 188pp; English.
XX
XX The invention describes an isolated nucleic acid encoding a CYP52A2B
CC protein comprising the fully defined sequence of 522 amino acids, as
CC given in the specification, and comprising a coding region defined by
CC nucleotides 1072-2640 of a fully defined sequence of 3755 base pairs, as
CC given in the specification. The nucleic acids encoding the cytochrome
CC P450 and NADH reductase enzymes of Candida tropicalis are useful for
CC increasing the production of dicarboxylic acid for oxidising organic
CC substrates such as fatty acids. This sequence represents a Candida
CC tropicalis associated polypeptide.
XX
XX Sequence 267 AA;
SQ

Query Match 86.4%; Score 19; DB 7; Length 267;
Best Local Similarity 57.1%; Pred. No. 1.8e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXXAEEF 7
Db 39 VTTTAEF 45

```

RESULT 10
ADFL1815
ID   ADF11815 standard; protein; 267 AA.
XX
XX   AC   ADF11815;
XX
XX   DT   12-FEB-2004 (first entry)
XX
XX   DE   C. tropicalis URA3A.
XX
XX   KW   CPRA; CPRB; CYP52A1A; CYP52A2A; CYP52A2B; CYP52A3A; CYP52A3B; CYP52A5A;
XX   KW   CYP52A5B; CYP52A8A; CYP52A8B; CYP52D4A; gene family; URA3A.
XX
XX   OS   Candida tropicalis.
XX
XX   PN   US2003148486-A1.
XX
XX   PD   07-AUG-2003.
XX
XX   PF   03-MAY-2002; 2002US-00139296.
XX
XX   PR   01-MAY-1998; 98US-0083798P.
XX   PR   05-OCT-1998; 98US-0103099P.
XX   PR   10-MAR-1999; 99US-0123555P.
XX   PR   30-APR-1999; 99US-00302620.
XX   PR   12-OCT-2001; 2001US-00976800.
XX
XX   PA   (WILS/) WILSON C R.
XX   PA   (CRAF/) CRAFT D L.
XX   PA   (EIRI/) EIRICH L D.
XX   PA   (ESHO/) ESHOO M.
XX   PA   (MADD/) MADDURI K M.
XX   PA   (CORN/) CORNETT C A.
XX   PA   (BREN/) BRENNER A A.
XX   PA   (TANG/) TANG M.
XX   PA   (LOPE/) LOPER J C.
XX   PA   (GLEE/) GLEESON M.
XX
XX   PI   Wilson CR, Craft DL, Eirich LD, Eshoo M, Madduri KM, Cornett CA;
XX   PI   Brenner AA, Tang M, Loper JC, Gleeson M;
XX   DR   WPI; 2003-897719/82.
XX   DR   N-PSDB; ADF11814.
XX
XX   PT   New CPRA, CPRB, CYP52A1A, CYP52A2A, CYP52A2B, CYP52A3A, CYP52A3B,
XX   PT   CYP52A5A, CYP52A5B, CYP52A8A, CYP52A8B or CYP52D4A gene, useful for
XX   PT   increasing production of dicarboxylic acid.
XX
XX   PS   Disclosure; SEQ ID NO 106; 194pp; English.
XX
XX   CC   The invention relates to a new isolated nucleic acid which encodes a
XX   CC   CPRA, CPRB, CYP52A1A, CYP52A2A, CYP52A2B, CYP52A3A, CYP52A3B, CYP52A5A,
XX   CC   CYP52A5B, CYP52A8A, CYP52A8B or CYP52D4A protein. The nucleic acid is
XX   CC   useful for discriminating between members of a gene family by quantifying
XX   CC   the amount of mRNA in a sample. The present sequence represents the amino
XX   CC   acid sequence of a Candida tropicalis protein.
XX
XX   SQ   Sequence 267 AA;
XX
XX   Query Match      86.4%; Score 19; DB 7; Length 267;
XX   Best Local Similarity 57.1%; Pred. No. 1.8e+03;
XX   Matches      4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY   1 VXXXAEEF 7
DB   39 VTTTAEF 45

RESULT 11
ADFL1697
ID   ADF11697 standard; protein; 267 AA.
XX
XX   AC   ADF11697;
XX
XX   DT   29-JAN-2004 (first entry)
XX
XX   DE   Integration vector pURA2in URA3A gene encoded protein.
XX
XX   KW   Integration vector pURA2in URA3A gene encoded protein.
XX
XX   OS   Candida tropicalis.
XX
XX   PN   US2003153060-A1.
XX
XX   PD   14-AUG-2003.
XX
XX   PF   03-MAY-2002; 2002US-00139218.
XX
XX   PR   01-MAY-1998; 98US-0083798P.
XX   PR   05-OCT-1998; 98US-0103099P.
XX   PR   10-MAR-1999; 99US-0123555P.
XX   PR   30-APR-1999; 99US-00302620.
XX   PR   12-OCT-2001; 2001US-00976800.
XX
XX   PA   (WILS/) WILSON C R.
XX   PA   (CRAF/) CRAFT D L.
XX   PA   (EIRI/) EIRICH L D.
XX   PA   (ESHO/) ESHOO M.
XX   PA   (MADD/) MADDURI K M.
XX   PA   (CORN/) CORNETT C A.
XX   PA   (BREN/) BRENNER A A.
XX   PA   (TANG/) TANG M.
XX   PA   (LOPE/) LOPER J C.
XX   PA   (GLEE/) GLEESON M.
XX
XX   PI   Wilson CR, Craft DL, Eirich LD, Eshoo M, Madduri KM, Cornett CA;
XX   PI   Brenner AA, Tang M, Loper JC, Gleeson M;
XX   DR   WPI; 2003-897719/82.
XX   DR   N-PSDB; ADF11814.
XX
XX   PT   New CPRA, CPRB, CYP52A1A, CYP52A2A, CYP52A2B, CYP52A3A, CYP52A3B,
XX   PT   CYP52A5A, CYP52A5B, CYP52A8A, CYP52A8B or CYP52D4A gene, useful for
XX   PT   increasing production of dicarboxylic acid.
XX
XX   PS   Disclosure; SEQ ID NO 106; 194pp; English.
XX
XX   CC   The invention relates to a new isolated nucleic acid which encodes a
XX   CC   CPRA, CPRB, CYP52A1A, CYP52A2A, CYP52A2B, CYP52A3A, CYP52A3B, CYP52A5A,
XX   CC   CYP52A5B, CYP52A8A, CYP52A8B or CYP52D4A protein. The nucleic acid is
XX   CC   useful for discriminating between members of a gene family by quantifying
XX   CC   the amount of mRNA in a sample. The present sequence represents the amino
XX   CC   acid sequence of a Candida tropicalis protein.
XX
XX   SQ   Sequence 267 AA;
XX
XX   Query Match      86.4%; Score 19; DB 7; Length 267;
XX   Best Local Similarity 57.1%; Pred. No. 1.8e+03;
XX   Matches      4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY   1 VXXXAEEF 7
DB   39 VTTTAEF 45

RESULT 12
ADE64232
ID   ADE64232 standard; protein; 267 AA.
XX
XX   AC   ADE64232;
XX
XX   DT   29-JAN-2004 (first entry)
XX
XX   DE   Integration vector pURA2in URA3A gene encoded protein.
XX
XX   KW   Integration vector pURA2in URA3A gene encoded protein.
XX
XX   OS   Candida tropicalis.
XX
XX   PN   US2003148486-A1.
XX
XX   PD   07-AUG-2003.
XX
XX   PF   03-MAY-2002; 2002US-00139296.
XX
XX   PR   01-MAY-1998; 98US-0083798P.
XX   PR   05-OCT-1998; 98US-0103099P.
XX   PR   10-MAR-1999; 99US-0123555P.
XX   PR   30-APR-1999; 99US-00302620.
XX   PR   12-OCT-2001; 2001US-00976800.
XX
XX   PA   (WILS/) WILSON C R.
XX   PA   (CRAF/) CRAFT D L.
XX   PA   (EIRI/) EIRICH L D.
XX   PA   (ESHO/) ESHOO M.
XX   PA   (MADD/) MADDURI K M.
XX   PA   (CORN/) CORNETT C A.
XX   PA   (BREN/) BRENNER A A.
XX   PA   (TANG/) TANG M.
XX   PA   (LOPE/) LOPER J C.
XX   PA   (GLEE/) GLEESON M.
XX
XX   PI   Wilson CR, Craft DL, Eirich LD, Eshoo M, Madduri KM, Cornett CA;
XX   PI   Brenner AA, Tang M, Loper JC, Gleeson M;
XX   DR   WPI; 2003-897719/82.
XX   DR   N-PSDB; ADF11696.
XX
XX   PT   New CPRA, CPRB, CYP52A1A, CYP52A2A, CYP52A2B, CYP52A3A, CYP52A3B,
XX   PT   CYP52A5A, CYP52A5B, CYP52A8A, CYP52A8B or CYP52D4A gene, useful for
XX   PT   discriminating members of a gene family.
XX
XX   PS   Disclosure; SEQ ID NO 106; 196pp; English.
XX
XX   CC   The invention relates to a new isolated nucleic acid which encodes a
XX   CC   CPRA, CPRB, CYP52A1A, CYP52A2A, CYP52A2B, CYP52A3A, CYP52A3B, CYP52A5A,
XX   CC   CYP52A5B, CYP52A8A, CYP52A8B or CYP52D4A protein. The nucleic acid is
XX   CC   useful for discriminating between members of a gene family by quantifying
XX   CC   the amount of mRNA in a sample. The present sequence represents the amino
XX   CC   acid sequence of a Candida tropicalis protein.
XX
XX   SQ   Sequence 267 AA;
XX
XX   Query Match      86.4%; Score 19; DB 7; Length 267;
XX   Best Local Similarity 57.1%; Pred. No. 1.8e+03;
XX   Matches      4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY   1 VXXXAEEF 7
DB   39 VTTTAEF 45

```

KW Yeast; enzyme; NADPH reductase; CPR; cytochrome P450; CYP;
 KW omega-hydroxylase; dicarboxylic acid; URA3.

OS Candida tropicalis.

XX US2003068800-A1.

XX 10-APR-2003.

XX 03-MAY-2002; 2002US-00139905.

XX 01-MAY-1998; 98US-0083798P.

PR 05-OCT-1998; 98US-0103099P.

PR 10-MAR-1999; 99US-0123555P.

PR 30-APR-1999; 99US-00302620.

PR 12-OCT-2001; 2001US-00976800.

XX (WILS/) WILSON C R.

PA (CRAF/) CRAFT D L.

PA (EIRI/) EIRICH L D.

PA (ESHO/) ESHOO M.

PA (MADD/) MADDURI K M.

PA (CORN/) CORNETT C A.

PA (BRENN/) BRENNER A A.

PA (TANG/) TANG M.

PA (LOPE/) LOPER J C.

PA (GLEE/) GLEESON M.

XX Wilson CR, Craft DL, Eirich LD, Eshoo M, Madduri KM, Cornett CA,
 PI Brenner AA, Tang M, Loper JC, Gleeson M;

XX WPI; 2004-020205/02.

XX N-PSDB; ADE64231.

XX Novel isolated CPRA, CPRB, CYP52A1A, CYP52A2A, CYP52A2B, CYP52A3A,
 PT CYP52A3B, CYP52A5A, CYP52A5B, CYP52A8A, CYP52A8B or CYP52B4A protein,
 PT useful for increasing production of dicarboxylic acid in cells.

XX Example 15; SEQ ID NO 106; 195pp; English.

XX The invention relates to an isolated CPRA, CPRB, CYP52A1A, CYP52A2A,
 CC CYP52A2B, CYP52A3A, CYP52A3B, CYP52A5A, CYP52A5B, CYP52A8A, CYP52A8B or
 CC CYP52D4A protein (CYP - cytochrome P450, CPR - NADPH reductase) of the
 CC Candida tropicalis omega-hydroxylase complex. Also included are the
 CC nucleic acids encoding the CYP/CPR proteins (including their coding
 CC regions), a vector comprising the nucleotide acid, a host cell
 CC transfected or transformed with the vector, discriminating members of a
 CC gene family by quantifying the amount of target mRNA in a sample and
 CC increasing production of a dicarboxylic acid (comprising: providing a
 CC host cell having a naturally occurring CPR/CYP protein and culturing the
 CC host cell in media containing an organic substrate which upregulates the
 CC genes, to effect increased production of dicarboxylic acid). The CYP and
 CC CPR proteins, present in higher levels than normal is useful for
 CC increasing production of dicarboxylic acids. The present sequence
 CC represents a C. tropicalis URA3 protein, used to make the CYP/CPR
 CC integration vectors.

XX Sequence 267 AA;

XX Query Match 86.4%; Score 19; DB 8; Length 267;

XX Best Local Similarity 57.1%; Pred. No. 1.8e+03;

XX Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXXAEF 7

DB 39 VTTTAEF 45

RESULT 13

ADG73846

ID ADG73846 standard; protein; 267 AA.

XX AC ADG73846;

XX 11-MAR-2004 (first entry)

XX Candida tropicalis URA3 gene-encoded polypeptide.

XX URA3; selectable marker; enzyme; beta-oxidation; dicarboxylic acid.

XX Candida tropicalis.

XX WO2003100013-A2.

XX 04-DEC-2003.

XX 22-MAY-2003; 2003WO-US016453.

XX 23-MAY-2002; 2002US-0383332P.

XX 22-MAY-2003; 2003US-00444467.

XX (COGN-) COGNIS CORP.

XX Craft DL;

XX WPI; 2004-081907/08.

XX N-PSDB; ADG73845.

XX New non-reversible beta-oxidation blocked Candida tropicalis cell having
 PT a disrupted chromosomal POX4 and/or POX5 gene(s), useful for producing
 PT dicarboxylic acids used in the preparation of perfumes or macrolid
 PT antibiotics.

XX Example 2; SEQ ID NO 10; 38pp; English.

XX The present sequence is that of the polypeptide encoded by the URA3A
 CC selectable marker gene of Candida tropicalis. A claimed C. tropicalis
 CC cell has a disrupted chromosomal POX4 and/or POX5 gene in which a portion
 CC of the gene(s) is deleted through homologous recombination with a
 CC selectable marker gene, especially URA3. This blocks the beta-oxidation
 CC pathway of the cell. The size of the deletion of the chromosomal gene
 CC prevents the strain from reverting to wild-type activity. The cell is
 CC used in a claimed process for producing alpha,omega-dicarboxylic acid.

XX Sequence 267 AA;

XX Query Match 86.4%; Score 19; DB 8; Length 267;

XX Best Local Similarity 57.1%; Pred. No. 1.8e+03;

XX Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXXAEF 7

DB 39 VTTTAEF 45

RESULT 14

ADJ26614

ID ADJ26614 standard; protein; 267 AA.

XX AC ADJ26614;

XX 20-MAY-2004 (first entry)

XX Candida tropicalis URA3A protein.

XX CYP; CPRA protein; CPRB protein; dicarboxylic acid; plastic; adhesive;

XX fragrance; URA3A.

XX Candida tropicalis.

XX US2003186411-A1.

XX 02-OCT-2003.

XX 03-APR-2003; 2003US-00405660.

CC *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The invention is also
 CC useful for the identification of potential new targets for antibiotic
 CC development. The antisense nucleic acids can also be used to identify
 CC proteins used in proliferation, to express these proteins, and to obtain
 CC antibodies capable of binding to the expressed proteins. The proteins can
 CC be used to screen compounds in rational drug discovery programmes. The
 CC antisense nucleic acid sequence is also useful to screen for homologous
 CC nucleic acids which are required for cell proliferation in a wide variety
 CC of organisms. The present sequence represents an essential prokaryotic
 CC cellular proliferation protein. Note: The sequence data for this patent
 CC did not form part of the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 298 AA;

Query Match 86.4%; Score 19; DB 4; Length 298;
 Best Local Similarity 57.1%; Pred. No. 2e+03;
 Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXXAEF 7
 | |||
 Db 172 VSTTAEF 178

RESULT 17
 ABU28707
 ID ABU28707 standard; protein; 298 AA.
 XX
 AC ABU28707;
 DT 19-JUN-2003 (first entry)
 XX
 DE Protein encoded by Prokaryotic essential gene #14234.

XX Antisense; prokaryotic essential gene; cell proliferation; drug design.
 XX
 OS *Escherichia coli*.
 XX
 PN WO20027183-A2.
 XX
 PD 03-OCT-2002.

PF 21-MAR-2002; 2002WO-US009107.

XX
 PR 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.

XX (ELIT-) ELITRA PHARM INC.

XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX WPI; 2003-0299926/02.
 DR N-PSDB; ACA32577.

XX New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.

XX Claim 25; SEQ ID NO 56631; 1766pp; English.

XX The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding

CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 298 AA;

Query Match 86.4%; Score 19; DB 6; Length 298;
 Best Local Similarity 57.1%; Pred. No. 2e+03;
 Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXXAEF 7
 | |||
 Db 172 VSTTAEF 178

RESULT 18
 ABG17869
 ID ABG17869 standard; protein; 305 AA.
 XX
 AC ABG17869;
 XX
 DT 18-FEB-2002 (first entry)
 XX
 DE Novel human diagnostic protein #17860.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 XX
 OS *Homo sapiens*.

XX WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US008631.

XX PR 31-MAR-2000; 2000US-00540217.

XX PR 23-AUG-2000; 2000US-00649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

XX N-PSDB; AAS82056.

XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.

XX Claim 20; SEQ ID NO 48228; 103pp; English.

CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic
 CC amino acid sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 CC
 XX SQ Sequence 305 AA;

Query Match 86.4%; Score 19; DB 4; Length 305;
 Best Local Similarity 57.1%; Pred. No. 2e+03;
 Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 VXXAAEF 7
 | |||
 Db 217 VSTTAEF 223

RESULT 19
 AAW69244
 ID AAW69244 standard; protein; 309 AA.
 AC AAW69244;
 XX 21-OCT-1998 (first entry)
 DT
 DE Clone CK201_1 protein sequence.
 XX
 XX Secreted protein; nutritional source; cell proliferation activity;
 KW cell differentiation activity; immune stimulant; tissue growth activator;
 KW haematopoiesis regulator; anti-inflammatory; tumour invasion suppressor;
 KW tumour inhibitor; clone CK201_1.
 XX
 XX Homo sapiens.

Key Location/Qualifiers
 FT Misc-difference 152 /note="encoded by ARG"
 FT
 PN WO9825962-A2.
 XX 18-JUN-1998.
 PD
 XX 12-DEC-1997; 97WO-US023224.
 PF
 XX 13-DEC-1996; 96US-00766263.
 PR 11-DEC-1997; 97US-00989232.
 XX
 PA (GEMY) GENETICS INST INC.
 XX
 XX Jacobs K, Mccoy JM, Lavallie ER, Racie LA, Merberg D, Treacy M;
 PI Spaulding V, Agostino MJ;
 PI WPI; 1998-362424/31.
 DR N-PSDB; AAV44863.
 XX
 XX New isolated polynucleotides - obtained from human adult testis, human
 PT adult ovary, human adult brain and human adult heart cDNA libraries.

XX
 PS Claim 27; Page 74-75; 108pp; English.
 XX
 CC This sequence is secreted protein, encoded by a polynucleotide of the
 CC invention. The DNA was isolated from a human adult testis cDNA library,
 CC and is designated clone CK201_1. The DNA sequences and encoded
 CC polypeptides can be used as nutritional sources or supplements, or may
 CC exhibit e.g. cytokine and cell proliferation/differentiation activity,
 CC immune stimulating or suppressing activity, haematopoiesis regulating
 CC activity, receptor/ligand activity, anti-inflammatory activity,
 CC activin/inhibin activity, chemostatic/chemokinetic activity,
 CC cadherin/tumour invasion suppressor activity, tissue growth activity,
 CC tumour inhibition activity or other activities
 XX

SQ Sequence 309 AA;

Query Match 86.4%; Score 19; DB 2; Length 309;
 Best Local Similarity 57.1%; Pred. No. 2.1e+03;
 Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 VXXAAEF 7
 | |||
 Db 275 VSAAEF 281

RESULT 20
 AAB90724
 ID AAB90724 standard; protein; 309 AA.
 XX
 AC AAB90724;
 XX
 XX 07-JUN-2001 (first entry)
 DT
 DE Human CK201_1 protein sequence SEQ ID 145.

XX
 XX Human; secreted protein; nutrient; cytokine modulator; proliferation;
 KW differentiation; immune system modulator; tissue growth; chemotactic;
 KW haemostatic; thrombolytic; anti-inflammatory; tumour inhibition;
 KW haematopoiesis.

OS Homo sapiens.
 XX
 XX WO200119988-A1.
 PN
 XX 22-MAR-2001.
 PD
 XX 14-SEP-2000; 2000WO-US025135.
 PF
 PR 17-SEP-1999; 99US-00398829.
 XX
 XX (GEMY) GENETICS INST INC.
 PA
 XX Jacobs K, Mccoy JM, Lavallie ER, Collins-Racie LA, Evans C;
 PI Merberg D, Treacy M, Bowman MR, Spaulding V, Agostino MJ;
 PI WPI; 2001-244801/25.
 DR N-PSDB; AAF98460.

XX Isolated nucleic acids encoding polypeptides, useful for modulating e.g.
 PT cytokine and cell proliferation/differentiation activity, the immune
 PT system and hematopoiesis regulating activity.

PS Disclosure; Page 473-474; 557pp; English.

XX Human cDNA clones represented in AAF98374 - AAF98489 encode secreted
 CC proteins AAB90667 - AAB90750. The cDNA clones are isolated from various
 CC tissue types, and may be used in the prevention, treatment and diagnosis
 CC of diseases associated with inappropriate protein expression. The
 CC polypeptides and nucleic acids may be used as nutrients or to modulate
 CC cytokine and cell proliferation/differentiation activity and may also be
 CC involved in modulation of the immune system. The cDNA sequences,
 CC proteins, their agonists and/or antagonists exhibit haematopoiesis
 CC regulating activity; tissue growth activity; activin/inhibin activity;

CC chemotactic/chemokinetic activity; haemostatic and thrombolytic activity;
 CC receptor/ligand activity; anti-inflammatory activity; haematopoiesis
 CC activity; cadherin/tumour suppressor activity; and/or tumour inhibition
 CC activity. Included in the invention are probes represented in AAF98490 -
 CC AAF98572 which are specific for the cDNA clones encoding the secreted
 CC proteins
 XX
 SQ Sequence 309 AA;

Query Match 86.4%; Score 19; DB 4; Length 309;
 Best Local Similarity 57.1%; Pred. No. 2.1e+03;
 Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXXXAEF 7
 | |||
 Db 275 VSAAAEF 281

RESULT 21
 ABU44714
 ID ABU44714 standard; protein; 338 AA.
 XX
 AC ABU44714;
 XX
 DT 19-JUN-2003 (first entry)
 DE Protein encoded by Prokaryotic essential gene #30241.
 XX
 KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
 XX
 OS Salmonella paratyphi.
 XX
 PN W0200277183-A2.
 XX
 PD 03-OCT-2002.
 XX
 PF 21-MAR-2002; 2002WO-US009107.
 XX
 PR 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 23-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 XX
 PA (ELIT-) ELITRA PHARM INC.
 XX
 PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX
 DR WPI; 2003-029926/02.
 DR N-PSDB; ACA48584.
 XX
 XX New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX
 PS Claim 25; SEQ ID NO 72638; 1766pp; English.
 XX
 CC The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies

CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 338 AA;

Query Match 86.4%; Score 19; DB 6; Length 338;
 Best Local Similarity 57.1%; Pred. No. 2.3e+03;
 Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXXXAEF 7
 | |||
 Db 271 VSSSAEF 277

RESULT 22
 AAB93191
 ID AAB93191 standard; protein; 697 AA.
 XX
 AC AAB93191;
 XX
 DT 26-JUN-2001 (first entry)
 DE Human protein sequence SEQ ID NO:12146.
 XX
 KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.
 XX
 OS Homo sapiens.
 XX
 PN EP1074617-A2.
 XX
 PD 07-FEB-2001.
 XX
 PF 28-JUL-2000; 2000EP-00116126.
 XX
 PR 29-JUL-1999; 99JP-00248036.
 PR 27-AUG-1999; 99JP-00300253.
 PR 11-JAN-2000; 2000JP-00118776.
 PR 02-MAY-2000; 2000JP-00183767.
 PR 09-JUN-2000; 2000JP-00241899.
 XX
 PA (HELI-) HELIX RES INST.
 XX
 PI Ora T, Isegai T, Nishikawa T, Hayaehi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX
 DR WPI; 2001-318749/34.
 XX
 PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-
 PT length cDNAs defined in the specification, and for the detection and/or
 PT diagnosis of the abnormality of the proteins encoded by the full-length
 PT cDNAs.
 XX
 PS Claim 8; SEQ ID NO 12146; 2537pp + Sequence Listing; English.
 XX
 CC The present invention describes primer sets for synthesising 5602 full-
 CC length cDNAs defined in the specification. Where a primer set comprises:
 CC (a) an oligo-dr primer and an oligonucleotide complementary to the
 CC complementary strand of a polynucleotide which comprises one of the 5602
 CC nucleotide sequences defined in the specification, where the

CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in the
CC specification. The primer sets can be used in antisense therapy and in
CC gene therapy. The primers are useful for synthesising polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to AAH95893
CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC oligonucleotides, all of which are used in the exemplification of the
CC present invention
XX
SQ Sequence 697 AA;

Query Match 86.4%; Score 19; DB 4; Length 697;
Best Local Similarity 57.1%; Pred. No. 5.1e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXXAEEF 7
| |||
Db 662 VSAAAEF 668

RESULT 23
ADL90006
ID ADL90006 standard; protein; 957 AA.

XX AC ADL90006;

XX DT 20-MAY-2004 (first entry)

XX DE Gluconobacter oxydans glucose-6-phosphoric acid isomerase protein.

XX KW transaldolase activity; glucose-6-phosphate isomerase; NADH production;
KW target substance manufacture; enzyme;
KW glucose-6-phosphoric acid isomerase.

XX OS Gluconobacter oxydans.

XX XN JP2004024140-A.

XX PD 29-JAN-2004.

XX PF 26-JUN-2002; 2002JP-00186487.

XX PR 26-JUN-2002; 2002JP-00186487.

XX PA (AJIN) AJINOMOTO KK.

XX WPI; 2004-127093/13.

DR N-PSDB; ADL90004.

PT Novel protein having transaldolase activity or glucose-6-phosphate
PT isomerase activity, useful for producing a target substance e.g.,
PT xylitol.

XX Claim 1; SEQ ID NO 18; 89pp; Japanese.

XX PS The invention comprises the amino acid and coding sequences of
CC Gluconobacter oxydans proteins which possess transaldolase activity
CC and/or glucose-6-phosphate isomerase activity. The DNA and protein
CC sequences of the invention are involved in the production of NADH. The
CC DNA and protein sequences of the invention are useful for manufacturing a
CC target substance. The present amino acid sequence represents the
CC Gluconobacter oxydans glucose-6-phosphoric acid isomerase protein.

XX SQ Sequence 957 AA;

Query Match 86.4%; Score 19; DB 8; Length 957;
Best Local Similarity 57.1%; Pred. No. 7.2e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXXAEEF 7
| |||
Db 861 VATAAEF 867

RESULT 24

ABU47745

ID ABU47745 standard; protein; 1266 AA.

XX AC ABU47745;

XX DT 19-JUN-2003 (first entry)

XX DE Protein encoded by Prokaryotic essential gene #33272.

XX KW Antisense; prokaryotic essential gene; cell proliferation; drug design.

XX OS Salmomella typhi.

XX PN WC20027183-A2.

XX PD 03-OCT-2002.

XX PF 21-MAR-2002; 2002WO-US009107.

XX PR 21-MAR-2001; 2001US-00815242.

XX PR 06-SEP-2001; 2001US-00948993.

XX PR 25-OCT-2001; 2001US-0342923P.

XX PR 08-FEB-2002; 2002US-00072851.

XX PR 06-MAR-2002; 2002US-0362699P.

XX PA (ELIT-) ELITRA PHARM INC.

XX PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX WPI; 2003-029926/02.

DR N-PSDB; ACA51615.

PT New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.

XX Claim 25; SEQ ID NO 75669; 1766pp; English.

XX CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required

CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 1266 AA;
 Query Match 86.4%; Score 19; DB 6; Length 1266;
 Best Local Similarity 57.1%; Pred. No. 9.9e+03;
 Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 VXXAAEF 7
 | |||
 Db 895 VSSAAEF 901
 RESULT 25
 ID ADJ80131 standard; protein; 1450 AA.
 XX
 AC ADJ80131;
 XX
 DT 06-MAY-2004 (first entry)
 XX
 DE Novel human nucleic acid-associated protein #7.
 XX
 KW cytostatic; antiarteriosclerotic; cerebroprotective; antiparkinsonian;
 KW anticonvulsant; anti-HIV; antiallergic; antibacterial; virucide;
 KW gene therapy; nucleic acid-associated protein; cancer; atherosclerosis;
 KW stroke; Parkinson's disease; epilepsy; Cushing's syndrome; AIDS; allergy;
 KW microarray element; protein-protein interaction; drug-target interaction;
 KW gene expression; chromosomal mapping; diagnosis.
 XX
 OS Homo sapiens.
 XX
 PN WO2003038052-A2.
 XX
 PD 08-MAY-2003.
 XX
 PF 29-OCT-2002; 2002WO-US034846.
 XX
 PR 29-OCT-2001; 2001US-0348442P.
 PR 01-NOV-2001; 2001US-0335544P.
 PR 05-NOV-2001; 2001US-0337535P.
 PR 09-NOV-2001; 2001US-0344650P.
 PR 15-NOV-2001; 2001US-0334762P.
 XX
 PA (INCY-) INCYTE GENOMICS INC.
 XX
 PI Becha SD, Borowsky ML, Burford N, Chawla NK, Elliott VS;
 PI Emerling BM, Forsythe IJ, Gietzen KJ, Gorvath AE, Griffin JA;
 PI Hafalia AJA, Ison CH, Lal PG, Lee EA, Lee S, Lee SV, Marquis JP;
 PI Ramkumar J, Sprague WM, Swarnakar A, Tang YT, Warren BA, Yang J;
 PI Yue H, Zebardjian Y;
 XX
 WPI; 2003-430514/40.
 DR N-PSDB; ADJ80189.
 XX
 XX New human nucleic acid-associated protein (NAAP) and polynucleotide,
 PT useful for diagnosing, treating, and preventing disorders associated with
 PT aberrant expression of NAAP, e.g. cancer, AIDS, stroke or infection.
 XX
 PS Claim 1; SEQ ID NO 7; 443pp; English.
 XX
 CC The invention relates to novel human nucleic acid-associated proteins and
 CC genes encoding them, sequences that have at least 90-99 % identity to the
 CC sequences; or biologically active or immunogenic fragments of these. The
 CC polypeptides and polynucleotides are useful in diagnosing, treating and
 CC preventing disorders associated with aberrant expression of NAAP, such as

CC cell proliferative (e.g. cancer or atherosclerosis), neurological (e.g.
 CC stroke, Parkinson's disease or epilepsy), developmental (e.g. Cushing's
 CC syndrome), autoimmune/inflammatory (e.g. AIDS or allergies), or
 CC infections. These may also be used as elements on a microarray which may
 CC monitor or measure protein-protein interactions, drug-target
 CC interactions, and gene expression profiles. The polynucleotide may also
 CC be used in chromosomal mapping and in various diagnostic assays. These
 CC are also useful in assessing the effects of exogenous compounds on the
 CC expression of nucleic acids and amino acid sequences of NAAP, in
 CC facilitating drug discovery process, and in investigating the
 CC pathogenesis of diseases or medical conditions. This sequence corresponds
 CC to one of the proteins of the inventions.
 XX
 SQ Sequence 1450 AA;
 Query Match 86.4%; Score 19; DB 7; Length 1450;
 Best Local Similarity 57.1%; Pred. No. 1.1e+04;
 Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 VXXAAEF 7
 | |||
 Db 1415 VSAAEF 1421
 RESULT 26
 ID ABR64238 standard; protein; 1584 AA.
 XX
 AC ABR64238;
 XX
 DT 15-OCT-2003 (first entry)
 XX
 DE Angiogenesis protein BNO147.
 XX
 KW Cytostatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological;
 KW antipsoriatic; antiarteriosclerotic; cardiant; vasotropic; angiogenesis;
 KW gene therapy; vasculature; cancer; rheumatoid arthritis; psoriasis;
 KW diabetic retinopathy; cardiovascular disease; atherosclerosis;
 KW ischemic limb disease; coronary artery disease.
 XX
 OS Homo sapiens.
 XX
 PN WO2003027285-A1.
 XX
 PD 03-APR-2003.
 XX
 PF 19-SEP-2002; 2002WO-AU001282.
 XX
 PR 27-SEP-2001; 2001AU-00007973.
 PR 27-SEP-2001; 2001AU-00007974.
 PR 11-OCT-2001; 2001AU-00008210.
 PR 29-OCT-2001; 2001AU-00008532.
 PR 13-NOV-2001; 2001AU-00008838.
 PR 28-AUG-2002; 2002AU-00951032.
 XX
 PA (BION-) BIONOMICS LTD.
 XX
 PI Gamble JR, Hahn CN, Vadas MA;
 XX
 WPI; 2003-354655/33.
 DR N-PSDB; ACF34513.
 XX
 XX New angiogenic genes and polypeptides, useful for diagnosing,
 PT prognosticating or treating an angiogenesis-related disorder, e.g.
 PT cancer, rheumatoid arthritis, diabetic retinopathy, psoriasis or
 PT cardiovascular diseases.
 XX
 PS Claim 15; SEQ ID NO 173; 90pp; English.
 XX
 CC The invention relates to the isolation of novel genes (ACF34446-ACF34559)
 CC encoding proteins (ABR64180-ABR64281) involved in the process of
 CC angiogenesis. The nucleic acid molecules are useful in identifying and/or
 CC obtaining full-length human genes involved in an angiogenic process. The

CC nucleic acid molecule, polypeptides or complexes encoded, cells or
 CC genetically modified non-human animals derived from these are useful for
 CC the screening of candidate pharmaceutical compounds used in treating
 CC angiogenesis-related disorders. They are also useful for diagnosing,
 CC prognosticating or treating an angiogenesis-related disorder, which
 CC involves uncontrolled or enhanced angiogenesis or is a disorder in which
 CC a decreased vasculature is of benefit (e.g. cancer, rheumatoid arthritis,
 CC diabetic retinopathy, psoriasis or cardiovascular diseases such as
 CC atherosclerosis), or involves inappropriately arrested or decreased
 CC angiogenesis or is a disorder in which an expanding vasculature is of
 CC benefit (e.g. ischemic limb disease or coronary artery disease). The
 CC modulator of expression or activity of the polypeptide encoded by the
 CC nucleic acid sequence is useful for manufacturing a medicament for the
 CC treatment of an angiogenesis-related disorder. This sequence corresponds
 CC to one of the novel angiogenic protein
 XX
 SQ Sequence 1584 AA;
 Query Match 86.4%; Score 19; DB 6; Length 1584;
 Best Local Similarity 57.1%; Pred. No. 1.3e+04;
 Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 VXXXAEEF 7
 | |||
 Db 1549 VSAAAEF 1555
 RESULT 27
 ADN24084
 ID ADN24084 standard; protein; 2214 AA.
 XX
 AC ADN24084;
 XX
 DT 02-DEC-2004 (first entry)
 XX
 DE Bacterial polypeptide #6737.
 XX
 KW Recombinant DNA construct; transformed plant; improved plant property;
 KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
 KW pathogen tolerance; pest tolerance; plant disease resistance;
 KW cell cycle pathway modification; plant growth regulator;
 KW homologous recombination; seed oil yield; protein yield; carbohydrate;
 KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
 KW bacterial polypeptide.
 XX
 OS Bacteria.
 XX
 PN US2003233675-A1.
 XX
 PD 18-DEC-2003.
 XX
 PF 20-FEB-2003; 2003US-00369493.
 XX
 PR 21-FEB-2002; 2002US-0360039P.
 XX
 PA (CAOY/) CAO Y.
 PA (HINK/) HINKLE G J.
 PA (SLAT/) SLATER S C.
 PA (CHEN/) CHEN X.
 PA (GOLD/) GOLDMAN B S.
 XX
 PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
 XX WPI; 2004-061375/06.
 DR
 XX New recombinant DNA construct comprising a promoter positioned to provide
 PT for expression of a polynucleotide encoding a polypeptide from a
 PT microbial source, useful for producing plants with improved properties.
 XX Claim 1; SEQ ID NO 6737; 122pp; English.
 PS
 XX The invention relates to a recombinant DNA construct comprising a
 CC promoter functional in a plant cell, where the promoter is positioned to

CC provide for expression of a polynucleotide encoding a polypeptide from a
 CC microbial source. The invention also relates to a transformed plant
 CC comprising the recombinant DNA construct and a method of producing a
 CC transformed plant having an improved property. The plant is a crop plant
 CC such as maize or soybean. The method of producing a transformed plant
 CC having an improved property comprises transforming a plant with the
 CC recombinant DNA construct and growing the transformed plant, where the
 CC polynucleotide or polypeptide is useful for improving plant properties.
 CC The recombinant DNA construct is useful for producing plants with
 CC improved plant properties, e.g. improved cold, heat or drought tolerance,
 CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
 CC increased resistance to plant disease, better growth rate by modification
 CC of the cell cycle pathway with plant growth regulators, increased rate of
 CC homologous recombination, modified seed oil or protein yield and/or
 CC content, improved yield by modification of carbohydrate, nitrogen or
 CC phosphorus use and/or uptake, by modification of photosynthesis or by
 CC providing improved plant growth and development under at least one stress
 CC condition, improved lignin production or improved galactomannan
 CC production. This sequence represents a bacterial polypeptide used in the
 CC scope of the invention. Note: The sequence data for this patent did not
 CC form part of the printed specification but was obtained in electronic
 CC format from USPTO at seqdata.uspto.gov/sequence.html.
 XX
 SQ Sequence 2214 AA;
 Query Match 86.4%; Score 19; DB 8; Length 2214;
 Best Local Similarity 57.1%; Pred. No. 1.8e+04;
 Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 VXXXAEEF 7
 | |||
 Db 1770 VTTSAEF 1776
 RESULT 28
 ABG78424
 ID ABG78424 standard; peptide; 12 AA.
 XX
 AC ABG78424;
 XX
 DT 15-NOV-2002 (first entry)
 XX
 DE Memapsin 2 substrate specificity determination peptide #30.
 XX
 KW Human; memapsin 2; aspartic protease; beta secretase;
 KW degenerative disease; Alzheimer's disease; amyloid precursor protein;
 KW APP; neuroprotective; neurotropic; inhibitor;
 KW substrate side-chain preference.
 XX
 OS Homo sapiens.
 OS Synthetic.
 XX
 PN WO200253594-A2.
 XX
 PD 11-JUL-2002.
 XX
 PF 28-DEC-2001; 2001WO-US050826.
 XX
 PR 28-DEC-2000; 2000US-0258705P.
 PR 14-MAR-2001; 2001US-0275756P.
 XX
 PA (OKLA-) OKLAHOMA MEDICAL RES FOUND.
 PA (UNII) UNIV ILLINOIS FOUND.
 XX
 PI Tang JJN, Koelsch G, Ghosh AK;
 XX WPI; 2002-619088/66.
 DR
 XX New memapsin 2 activity inhibitor useful in treatment of e.g. Alzheimer's
 PT disease.
 PT
 XX Example 5; Page 52; 74pp; English.
 PS
 XX

CC The invention relates to an inhibitor of catalytically active memapsin 2
 CC (an aspartic protease which can cleave at beta secretase sites), which
 CC binds to the active site of memapsin 2 defined by the presence of two
 CC catalytic aspartic residues and substrate binding cleft. Also included is
 CC a method of determination of the substrate side-chain preference in
 CC memapsin 2 sub-sites comprising: (a) reacting a mixture of memapsin 2
 CC substrates with memapsin 2, and determining the sub-site preference of
 CC memapsin 2 by determining relative initial hydrolysis rates of the
 CC mixture of memapsin 2 substrates; or (b) preparing a combinatorial
 CC library of memapsin 2 inhibitors containing a base sequence taken from
 CC OH99-2 (Glu-Val-Asn-Leu-Ala-Glu-phe), probing the library of
 CC inhibitors with memapsin 2 which binds to several inhibitors to generate
 CC several bound memapsin 2, and detecting the bound memapsin 2 with an
 CC antibody raised to memapsin 2 and an alkaline phosphatase conjugated
 CC secondary antibody. The inhibitors may be used in the manufacture of a
 CC medicament for the treatment of Alzheimer's disease since memapsin 2 may
 CC be involved in the cleavage of amyloid precursor protein (APP), and for
 CC determining the substrate side-chain preference in memapsin 2 sub-sites.
 CC The present sequence represents a subsite variant peptide used to
 CC determine the substrate specificity of human memapsin 2
 CC
 XX
 SQ Sequence 12 AA;

Query Match 81.8%; Score 18; DB 5; Length 12;
 Best Local Similarity 57.1%; Pred. No. 1.2e+02;
 Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXXAEEF 7
 | | | |
 DB 6 VNAAAEF 12

RESULT 29
 ABG78419
 ID ABG78419 standard; peptide; 12 AA.
 AC
 XX
 XX
 DT 15-NOV-2002 (first entry)
 DE Memapsin 2 substrate specificity determination peptide #25.
 XX Human; memapsin 2; aspartic protease; beta secretase;
 KW degenerative disease; Alzheimer's disease; amyloid precursor protein;
 KW APP; neuroprotective; nootropic; inhibitor;
 KW substrate side-chain preference.
 XX
 OS Homo sapiens.
 OS Synthetic.
 XX
 XX WO200253594-A2.
 XX
 XX PD 11-JUL-2002.
 XX
 XX 28-DEC-2001; 2001WO-US050826.
 XX
 XX 28-DEC-2000; 2000US-0258705P.
 PR 14-MAR-2001; 2001US-0275756P.
 XX
 XX (OKLA-) OKLAHOMA MEDICAL RES FOUND.
 PA (UNII) UNIV ILLINOIS FOUND.
 XX
 XX Tang JUN, Koelsch G, Ghosh AK;
 XX
 XX WPI; 2002-619088/66.
 DR
 XX New memapsin 2 activity inhibitor useful in treatment of e.g. Alzheimer's
 XX disease.
 PT
 XX Example 5; Page 52; 74pp; English.
 XX
 PS The invention relates to an inhibitor of catalytically active memapsin 2
 CC (an aspartic protease which can cleave at beta secretase sites), which
 CC binds to the active site of memapsin 2 defined by the presence of two
 CC catalytic aspartic residues and substrate binding cleft. Also included is
 CC a method of determination of the substrate side-chain preference in
 CC memapsin 2 sub-sites comprising: (a) reacting a mixture of memapsin 2
 CC substrates with memapsin 2, and determining the sub-site preference of
 CC memapsin 2 by determining relative initial hydrolysis rates of the
 CC mixture of memapsin 2 substrates; or (b) preparing a combinatorial
 CC library of memapsin 2 inhibitors containing a base sequence taken from
 CC OH99-2 (Glu-Val-Asn-Leu-Ala-Glu-phe), probing the library of
 CC inhibitors with memapsin 2 which binds to several inhibitors to generate
 CC several bound memapsin 2, and detecting the bound memapsin 2 with an
 CC antibody raised to memapsin 2 and an alkaline phosphatase conjugated
 CC secondary antibody. The inhibitors may be used in the manufacture of a
 CC medicament for the treatment of Alzheimer's disease since memapsin 2 may
 CC be involved in the cleavage of amyloid precursor protein (APP), and for
 CC determining the substrate side-chain preference in memapsin 2 sub-sites.
 CC The present sequence represents a subsite variant peptide used to
 CC determine the substrate specificity of human memapsin 2
 CC
 XX
 SQ Sequence 12 AA;

CC binds to the active site of memapsin 2 defined by the presence of two
 CC catalytic aspartic residues and substrate binding cleft. Also included is
 CC a method of determination of the substrate side-chain preference in
 CC memapsin 2 sub-sites comprising: (a) reacting a mixture of memapsin 2
 CC substrates with memapsin 2, and determining the sub-site preference of
 CC memapsin 2 by determining relative initial hydrolysis rates of the
 CC mixture of memapsin 2 substrates; or (b) preparing a combinatorial
 CC library of memapsin 2 inhibitors containing a base sequence taken from
 CC OH99-2 (Glu-Val-Asn-Leu-Ala-Glu-phe), probing the library of
 CC inhibitors with memapsin 2 which binds to several inhibitors to generate
 CC several bound memapsin 2, and detecting the bound memapsin 2 with an
 CC antibody raised to memapsin 2 and an alkaline phosphatase conjugated
 CC secondary antibody. The inhibitors may be used in the manufacture of a
 CC medicament for the treatment of Alzheimer's disease since memapsin 2 may
 CC be involved in the cleavage of amyloid precursor protein (APP), and for
 CC determining the substrate side-chain preference in memapsin 2 sub-sites.
 CC The present sequence represents a subsite variant peptide used to
 CC determine the substrate specificity of human memapsin 2
 CC
 XX
 SQ Sequence 12 AA;

Query Match 81.8%; Score 18; DB 5; Length 12;
 Best Local Similarity 57.1%; Pred. No. 1.2e+02;
 Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXXAEEF 7
 | | | |
 DB 6 VNTAAEF 12

RESULT 30
 ABG78418
 ID ABG78418 standard; peptide; 12 AA.
 AC
 XX
 XX
 DT 15-NOV-2002 (first entry)
 DE Memapsin 2 substrate specificity determination peptide #24.
 XX Human; memapsin 2; aspartic protease; beta secretase;
 KW degenerative disease; Alzheimer's disease; amyloid precursor protein;
 KW APP; neuroprotective; nootropic; inhibitor;
 KW substrate side-chain preference.
 XX
 OS Homo sapiens.
 OS Synthetic.
 XX
 XX WO200253594-A2.
 XX
 XX PD 11-JUL-2002.
 XX
 XX 28-DEC-2001; 2001WO-US050826.
 XX
 XX 28-DEC-2000; 2000US-0258705P.
 PR 14-MAR-2001; 2001US-0275756P.
 XX
 XX (OKLA-) OKLAHOMA MEDICAL RES FOUND.
 PA (UNII) UNIV ILLINOIS FOUND.
 XX
 XX Tang JUN, Koelsch G, Ghosh AK;
 XX
 XX WPI; 2002-619088/66.
 DR
 XX New memapsin 2 activity inhibitor useful in treatment of e.g. Alzheimer's
 XX disease.
 PT
 XX Example 5; Page 52; 74pp; English.
 XX
 PS The invention relates to an inhibitor of catalytically active memapsin 2
 CC (an aspartic protease which can cleave at beta secretase sites), which
 CC binds to the active site of memapsin 2 defined by the presence of two
 CC catalytic aspartic residues and substrate binding cleft. Also included is
 CC a method of determination of the substrate side-chain preference in
 CC memapsin 2 sub-sites comprising: (a) reacting a mixture of memapsin 2
 CC substrates with memapsin 2, and determining the sub-site preference of
 CC memapsin 2 by determining relative initial hydrolysis rates of the
 CC mixture of memapsin 2 substrates; or (b) preparing a combinatorial
 CC library of memapsin 2 inhibitors containing a base sequence taken from
 CC OH99-2 (Glu-Val-Asn-Leu-Ala-Glu-phe), probing the library of
 CC inhibitors with memapsin 2 which binds to several inhibitors to generate
 CC several bound memapsin 2, and detecting the bound memapsin 2 with an
 CC antibody raised to memapsin 2 and an alkaline phosphatase conjugated
 CC secondary antibody. The inhibitors may be used in the manufacture of a
 CC medicament for the treatment of Alzheimer's disease since memapsin 2 may
 CC be involved in the cleavage of amyloid precursor protein (APP), and for
 CC determining the substrate side-chain preference in memapsin 2 sub-sites.
 CC The present sequence represents a subsite variant peptide used to
 CC determine the substrate specificity of human memapsin 2
 CC
 XX
 SQ Sequence 12 AA;

CC a method of determination of the substrate side-chain preference in
CC memapsin 2 sub-sites comprising: (a) reacting a mixture of memapsin 2
CC substrates with memapsin 2, and determining the sub-site preference of
CC memapsin 2 by determining relative initial hydrolysis rates of the
CC mixture of memapsin 2 substrates; or (b) preparing a combinatorial
CC library of memapsin 2 inhibitors containing a base sequence taken from
CC OM99-2 (Glu-Val-Asn-Leu-Ala-Glu-phe), probing the library of
CC inhibitors with memapsin 2 which binds to several inhibitors to generate
CC several bound memapsin 2, and detecting the bound memapsin 2 with an
CC antibody raised to memapsin 2 and an alkaline phosphatase conjugated
CC secondary antibody. The inhibitors may be used in the manufacture of a
CC medicament for the treatment of Alzheimer's disease since memapsin 2 may
CC be involved in the cleavage of amyloid precursor protein (APP), and for
CC determining the substrate side-chain preference in memapsin 2 sub-sites.
CC The present sequence represents a subsite variant peptide used to
CC determine the substrate specificity of human memapsin 2
XX
SQ Sequence 12 AA;

Query Match 81.8%; Score 18; DB 5; Length 12;
Best Local Similarity 57.1%; Pred. No. 1.2e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 VXXXAEEF 7
DB 6 VNSAAEF 12
| |||
| |||

RESULT 31
AAP82869
ID AAP82869 standard; protein; 13 AA.
XX
AC AAP82869;
XX
DT 25-MAR-2003 (revised)
DT 12-NOV-1990 (first entry)
XX
DE OMP A-2 signal peptide/heterologous protein junction.
XX
KW lipoprotein; E.coli outer membrane protein (OMP) A signal peptide.
XX
OS Escherichia coli.
XX
FH Key Location/Qualifiers
FT Cleavage-site 4..5
FT /label= signal peptide/ protein junction
XX
PN US4757013-A.
XX
PD 12-JUL-1988.
XX
PF 26-JUN-1986; 86US-00880358.
XX
PR 25-JUL-1983; 83US-00494040.
PR 04-MAY-1984; 84US-00607224.
XX
PA (UYNV-) UNIV OF NEW YORK.
XX
PI Inouye M, Masui Y;
XX
DR WPI; 1988-212802/30.
DR N-PSDB; AAN80125.
XX
PT Cloning vehicles for polypeptide expression in bacterial hosts - has DNA
PT sequence coding for the signal peptide of OMP A to direct extracellular
PT expression.
XX
PS Disclosure; Page ?; 70pp; English.
XX
CC Efficiency of expression of lipoprotein expression vectors is improved by
CC introducing the omp A signal peptide from E.coli in place of lipoprotein
CC signal sequence. A fusion protein is produced where the desired protein
CC carries the omp A signal peptide at its N-terminal end. This signal

CC directs translocation of the product across the cytoplasmic membrane and
CC is then cleaved off. No modification of the cleavage site is necessary
CC during the secretion process, unlike for the lipoprotein signal peptide.
CC Synthetic linkers were inserted to ensure the structural gene was in
CC phase with the signal sequence. Proteins encoded by plasmids omp A-1 and
CC omp A-3 demonstrate the other two reading frames. (AAN80124 and
CC AAN80126). See also AAN80118-9 and AAN80128. (Updated on 25-MAR-2003 to
CC correct PA field.)
XX
SQ Sequence 13 AA;

Query Match 81.8%; Score 18; DB 1; Length 13;
Best Local Similarity 57.1%; Pred. No. 1.3e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 VXXXAEEF 7
DB 1 VAQAEEF 7
| |||
| |||

RESULT 32
ADO52379
ID ADO52379 standard; peptide; 41 AA.
XX
AC ADO52379;
XX
DT 12-AUG-2004 (first entry)
XX
DE Human ubiquitin peptide.
XX
KW Human; ubiquitin; annotated sequence; query sequence; database;
KW secondary structure characteristic; bio-dictionary.
XX
OS Homo sapiens.
XX
PN US2004101903-A1.
XX
PD 27-MAY-2004.
XX
PF 27-NOV-2002; 2002US-00305582.
XX
PR 27-NOV-2002; 2002US-00305582.
XX
PA (IBMC) INT BUSINESS MACHINES CORP.
XX
PI Rigoutsos I;
XX
DR WPI; 2004-459887/43.
XX
PT Annotating a query sequence e.g. protein sequence involves accessing
PT patterns associated with a database comprising annotated sequences,
PT assigning attributes to the patterns based on the annotated sequence and
PT analyzing the query sequence.
XX
PS Example 1; Fig 4; 28pp; English.

CC The invention relates to annotating a query sequence involving accessing
CC patterns associated with a database comprising annotated sequences,
CC assigning attributes to the patterns based on the annotated sequence and
CC using the patterns with assigned attributes to analyse the query
CC sequence. Also included are an apparatus for annotating a query
CC and an article of manufacture for annotating a query sequence. The method
CC further comprises the step of selecting the accessed patterns that match
CC the query sequence, storing the patterns with assigned attributes in a
CC database and defining an attribute vector from the patterns with assigned
CC attributes, the attribute vector characterising portions of the query
CC sequence. The query sequence is a polypeptide sequence comprising amino
CC acid residues and the attribute vector comprises a number of counters
CC (the number of counters is proportional to the number of amino acid
CC residues in the query sequence). The assigned attributes are used to
CC contribute values to counters of the attribute vector corresponding to
CC portions of the query sequence matched by the patterns. The method also
CC comprises a plurality of attribute vectors where the plurality of

CC attribute vectors are ranked, and the top ranking attribute vectors are
 CC reported. The method comprises the step of determining a score for the
 CC patterns with assigned attributes used to contribute to the attribute
 CC vector. The score represents a degree of similarity between the query
 CC sequence and the annotated sequences of the database and the score is
 CC normalised. The attributes relate to at least one of secondary structure
 CC characteristics of the query, presence of known domains, signal peptides,
 CC active sites, post-translationally modified sites, cytoplasmic behaviour,
 CC extracellular behaviour, and similarity of the query to each of the three
 CC phylogenetic domains as a function of amino acid position. The method is
 CC used for analysing or annotating query sequence e.g. protein or amino
 CC acid sequence, for elucidating properties of protein (including
 CC functional, structural and physicochemical properties) directly from
 CC corresponding amino acid sequence and for forming a bio-dictionary. The
 CC method is also useful for determining cytoplasmic, transmembrane or
 CC extracellular region of a query sequence and allows automated elucidation
 CC for annotating query sequences or complete genomes, quickly, exhaustively
 CC and objectively and avoids tedious manual analysis as reported
 CC previously. The present sequence is a peptide from human ubiquitin,
 CC analysed using the method of the invention.

XX Sequence 41 AA;

Query Match 81.8%; Score 18; DB 8; Length 41;
 Best Local Similarity 57.1%; Pred. No. 4.5e+02;
 Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VXXAAEF 7
 | | | |
 Db 25 VAALAEF 31

RESULT 33

AA861451
 ID AAB61451 standard; protein; 44 AA.

XX AC AAB61451;

XX 04-APR-2001 (first entry)

XX Human TANGO 275 TB domain #3.

XX TANGO; MANGO; colon; inflammation; tumor; renal; liver; lung; autoimmune;
 KW allergy; cardiovascular; brain; degenerative; placental; pancreatic;
 KW skeletal; muscle.

XX Homo sapiens.

XX WO200100672-A1.

XX 04-JAN-2001.

XX 29-JUN-2000; 2000WO-US018184.

XX 29-JUN-1999; 99US-00342687.

XX (MILL-) MILLENNIUM PHARM INC.

XX Holtzman DA, Barnes TM, Fraser CC, Sharp JD;

XX WPI; 2001-050127/06.

XX Isolated cDNAs encoding TANGO 244, TANGO 246, TANGO 275, TANGO 300 and
 PT MANGO 245 proteins, useful in the treatment of inflammatory diseases
 PT (e.g. idiopathic ulcerative colitis), tumors, renal disorders and liver
 PT disorders (e.g. jaundice).

XX Disclosure; Fig 12; 262pp; English.

XX The present invention relates to cDNAs encoding TANGO 244, TANGO 246,
 CC TANGO 275, TANGO 300 and MANGO 245 proteins. The nucleic acids, proteins
 CC and protein modulators are useful for treating colonic disorders,
 CC inflammatory diseases, tumors, renal disorders, liver disorders, lung

CC disorders, autoimmune diseases, allergic diseases, cardiovascular
 CC diseases, brain disorders, degenerative diseases placental, pancreatic,
 CC skeletal and muscle disorders

XX Sequence 44 AA;

Query Match 81.8%; Score 18; DB 4; Length 44;
 Best Local Similarity 57.1%; Pred. No. 4.9e+02;
 Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VXXAAEF 7
 | | | |
 Db 35 VYSSAEF 41

RESULT 34

AA47175
 ID AAM47175 standard; peptide; 47 AA.

XX AC AAM47175;

XX 12-FEB-2002 (first entry)

XX Modular enzyme system related ACP-domain N-terminal peptide MTAD.

XX Modular enzyme system; cyclic gene synthesis; repetitive coding sequence;
 KW antibiotic; non-ribosomal peptide synthetase; NRPS; PKS;
 KW polyketide synthase; actinomycin biosynthesis.

XX Stigmatella aurantiaca.

XX WO200181564-A2.

XX 01-NOV-2001.

XX 25-APR-2001; 2001WO-DE001578.

XX 26-APR-2000; 2000DE-01021267.

XX (ACTI-) ACTINODRUG PHARM GMBH.

XX Schauwecker F;

XX WPI; 2002-049276/06.

XX Preparing DNA encoding modular protein for e.g. producing new enzymes for
 PT synthesis of polyketide antibiotics, comprises cyclic integration of
 PT fragments into a vector.

XX Example 1; Fig 9; 83pp; German.

XX The present invention relates to the preparation of DNA, in a circular
 CC vector, that encodes one or more segments of a modular polypeptide. DNA
 CC or DNA libraries produced this way are used to produce modular
 CC polypeptides, particularly enzymes, which can be used to act on
 CC substrates to produce compounds for therapeutic testing. Enzymes of
 CC particular interest are those involved in non-ribosomal peptide synthesis
 CC or polyketide synthesis, and compounds for testing are particularly
 CC macrolide antibiotics, including penicillins, vancomycins or
 CC erythromycins, but may also be modular receptors. The present sequence is
 CC a peptide used in the exemplification of the invention

XX Sequence 47 AA;

Query Match 81.8%; Score 18; DB 5; Length 47;
 Best Local Similarity 57.1%; Pred. No. 5.3e+02;
 Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VXXAAEF 7
 | | | |
 Db 36 VSALAEF 42

```
RESULT 35
ADI21706
ID ADI21706 standard; protein; 53 AA.
XX
XX
AC ADI21706;
XX
XX
DT 15-APR-2004 (first entry)
XX
XX
DE Novel human polypeptide #185.
XX
XX
KW forensic; nutritional source; damaged tissue; diseased tissue;
KW myeloid cell disorder; lymphoid cell disorder;
KW bone cartilage tissue growth; tendon tissue growth;
KW ligament tissue growth; nerve tissue growth; regeneration; wound healing;
KW tissue repair; tissue replacement; burn; incision; ulcer; cancer; human.
XX
XX
OS Homo sapiens.
XX
XX
PN W02003025148-A2.
XX
XX
PD 27-MAR-2003.
XX
XX
PF 19-SEP-2002; 2002WO-US029964.
XX
XX
PR 19-SEP-2001; 2001US-0323739P.
PR 13-SEP-2002; 2002US-00323739.
XX
XX
PA (HYSE-) HYSEQ INC.
XX
XX
PI Tang YT, Asundi V, Goodrich RW, Ren F, Zhang J, Zhao QA, Wang J;
PI Ghosh M, Xue AJ, Wehrman T, Weng G, Zhou P, Drmanac RT, Wang D;
PI Haley-Vicente D;
XX
XX
DR WPI; 2003-354603/33.
DR N-PSDB; ADI21486.
XX
XX
PT New polynucleotides and secreted proteins, useful for treating myeloid or
PT lymphoid cell disorders, in bone cartilage, tendon, ligament and nerve
PT tissue growth or regeneration, in wound healing, and in tissue repair and
PT replacement.
XX
XX
PS Example 3; SEQ ID NO 957; 156pp; English.
XX
XX
CC The invention relates to an isolated polynucleotide encoding a
CC polypeptide with biological activity. The polynucleotides and
CC polypeptides are useful in diagnostics, forensics, gene mapping,
CC identification of mutations responsible for genetic disorders and other
CC traits, to assess biodiversity, as nutritional sources or supplements.
CC The polynucleotides may also be used as molecular weight markers,
CC chromosome markers or map related gene positions, or as an antigen to
CC raise anti-DNA antibodies or elicit immune response. The polypeptides are
CC useful for raising antibodies, as markers for tissues in which the
CC corresponding polypeptide is expressed, for re-engineering damaged or
CC diseased tissues, for treating myeloid or lymphoid cell disorders, in
CC bone cartilage, tendon, ligament and/or nerve tissue growth or
CC regeneration, in wound healing, in tissue repair and replacement, in
CC healing of burns, incisions and ulcers, and in treating cancer. The
CC present sequence represents the amino acid sequence of a novel human
CC polypeptide.
XX
XX
SQ Sequence 53 AA;
XX
XX
Query Match 81.8%; Score 18; DB 7; Length 53;
Best Local Similarity 57.1%; Pred. No. 6e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 VXXAAEF 7
DB 27 VRTSAEF 33
XX
XX
RESULT 36
AAY14477
ID ADI212304 standard; protein; 56 AA.
XX
XX
AC AAY14477;
XX
XX
DT 17-AUG-1999 (first entry)
XX
XX
DE Fragment of human secreted protein encoded by gene 7.
XX
XX
KW Human; secreted protein; fusion protein; gene therapy; protein therapy;
KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
KW developmental abnormality; foetal deficiency; blood; allergy; renal;
KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
KW inflammation; ischaemic shock; Alzheimer's disease; retinosis; AIDS;
KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
XX
XX
OS Homo sapiens.
XX
XX
PN W09919339-A1.
XX
XX
PD 22-APR-1999.
XX
XX
PF 08-OCT-1998; 98WO-US021142.
XX
XX
PR 09-OCT-1997; 97US-0061463P.
PR 09-OCT-1997; 97US-0061527P.
PR 09-OCT-1997; 97US-0061529P.
PR 09-OCT-1997; 97US-0061532P.
PR 09-OCT-1997; 97US-0061536P.
PR 09-OCT-1997; 97US-0071498P.
XX
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX
PI Brewer LA, Olsen HS, Duan R, Ebner R, Rosen CA, Ruben SM;
PI Florence KA, Young PE, Greene JM, Yu G, Ferrie AM, Florence C;
XX
XX
DR WPI; 1999-277587/23.
XX
XX
PT New isolated human genes and the secreted polypeptides they encode.
XX
XX
PS Disclosure; Page 14; 226pp; English.
XX
XX
CC This sequence represents a fragment of a secreted human protein encoded
CC by the nucleic acid molecule detailed in the descriptor line. The gene
CC can be used to generate fusion proteins by linking to the gene to a human
CC immunoglobulin Fc portion (e.g. AAX79002) for increasing the stability of
CC the fused protein as compared to the human protein only. The invention
CC relates to 53 novel genes and their fragments (nucleic acid sequences:
CC AAX79011-X79064; amino acid sequences AAY14411-Y14464) which are useful
CC for preventing, treating or ameliorating medical conditions e.g. by
CC protein or gene therapy. Also, pathological conditions can be diagnosed
CC by determining the amount of the new polypeptides in a sample or by
CC determining the presence of mutations in the new polynucleotides.
CC Specific uses are described for each of the 53 polynucleotides, based on
CC which tissues they are most highly expressed in (see AAX79011 for
CC described uses)
XX
XX
SQ Sequence 56 AA;
XX
XX
Query Match 81.8%; Score 18; DB 2; Length 56;
Best Local Similarity 57.1%; Pred. No. 6.4e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 VXXAAEF 7
DB 36 VAARAEF 42
XX
XX
RESULT 37
ADJ12304
ID ADJ12304 standard; protein; 56 AA.
XX
XX
```


AC ADJ12304;
 XX 20-MAY-2004 (first entry)
 DE Peptide feature of a human secreted protein SeqID 158.
 XX human; secreted; cancer; haematopoietic disease; anaemia;
 KW multiple myeloma; reproductive system disorder; prostatitis;
 KW inguinal hernia; musculoskeletal disease; systemic lupus erythematosus;
 KW gout; cardiovascular disease; arrhythmia; hypernatraemia; fetal disease;
 KW fetal alcohol syndrome; Down's syndrome; excretory disease;
 KW urinary incontinence; renal disorder; neural; sensory disease;
 KW Alzheimer's disease; meningitis; respiratory disease; emphysema;
 KW occupational lung disease; endocrine disease; diabetes;
 KW irritable bowel syndrome; epithelial disease; scleroderma;
 KW epidermolysis bullosa; cytostatic; antianemic; antiarrhythmic;
 KW antiasthmatic; anti-HIV; immunosuppressive; antiinflammatory;
 KW antiparkinsonian; tranquilizer; osteopathic; dermatological; antigout;
 KW immunomodulator; antiarrhythmic; cardiant; nootropic; antilipemic;
 KW nephrotropic; uropathic; neuroprotective; antiparkinsonian; tranquilizer;
 KW antidiabetic; anabolic; hypertensive; vulnerary.
 XX Homo sapiens.
 OS
 XX US2004010132-A1.
 PN
 XX 15-JAN-2004.
 PD
 XX 30-OCT-2001; 2001US-00984429.
 PF
 XX 09-OCT-1997; 97US-0061463P.
 PR 09-OCT-1997; 97US-0061527P.
 PR 09-OCT-1997; 97US-0061529P.
 PR 09-OCT-1997; 97US-0061532P.
 PR 09-OCT-1997; 97US-0061536P.
 PR 09-OCT-1997; 97US-0071498P.
 PR 08-OCT-1998; 98WO-US021142.
 PR 08-APR-1999; 99US-00288143.
 PR 01-NOV-2000; 2000US-0244591P.
 XX (ROSE/) ROSEN C A.
 PA (BREW/) BREWER L A.
 PA (DUAN/) DUAN R D.
 PA (RUBE/) RUBEN S M.
 PA (FLOR/) FLORENCE K A.
 PA (GREE/) GREENE J M.
 PA (YOUN/) YOUNG P E.
 PA (FERR/) FERRIE A M.
 PA (YUGG/) YU G.
 PA (FLOR/) FLORENCE C.
 PA (EBNE/) EBNER R.
 PA (OLSE/) OLSEN H.
 XX
 PI Rosen CA, Brewer LA, Duan RD, Ruben SM, Florence KA, Greene JM;
 PI Young PE, Ferrie AM, Yu G, Florence C, Ebner R, Olsen H;
 XX WPI; 2004-090518/09.
 XX New isolated nucleic acids and polypeptides, useful for diagnosing,
 PT treating, preventing or ameliorating diseases or disorders e.g. cancer,
 PT anemia, arthritis, asthma, inflammatory bowel disease or Alzheimer's
 PT disease.
 XX Disclosure; SEQ ID NO 158; 286pp; English.
 XX This invention relates to novel polynucleotides encoding human secreted
 CC proteins. Specifically, it refers to the vectors, host cells, recombinant
 CC and synthetic methods for producing human polynucleotides, polypeptides
 CC and antibodies. Furthermore, it relates to screening methods to identify
 CC agonists and antagonists that can be used to inhibit or enhance the
 CC production and function of the secreted proteins. The present invention
 CC describes these compositions as useful for diagnosing, treating or

CC preventing disorders such as cancer, haematopoietic diseases including
 CC anaemia and multiple myeloma, reproductive system disorders including
 CC prostatitis and inguinal hernia, musculoskeletal diseases including
 CC systemic lupus erythematosus and gout, cardiovascular disease including
 CC arrhythmia and hypernatraemia, mixed fetal diseases including fetal
 CC alcohol syndrome and Down's syndrome, excretory diseases including
 CC urinary incontinence and renal disorders, neural or sensory disease
 CC including Alzheimer's disease and meningitis, respiratory disease
 CC including emphysema and occupational lung disease, endocrine diseases
 CC including diabetes and glomerulonephritis, digestive diseases including
 CC portal hypertension and irritable bowel syndrome and connective tissue or
 CC epithelial diseases including scleroderma and epidermolysis bullosa. As
 CC such, there are various activities such as cytostatic, antianemic,
 CC antiarrhythmic, antiasthmatic, anti-HIV, immunosuppressive,
 CC dermatological, antiparkinsonian, antiinflammatory, cardiant, osteopathic,
 CC nootropic, antilipemic, nephrotropic, uropathic, antiarrhythmic, cardiant,
 CC antiparkinsonian, tranquilizer, osteopathic, anabolic, hypertensive and
 CC vulnerary. This polypeptide is a feature of a human secreted protein of
 CC the invention. NOTE: This sequence does not appear in the printed
 CC specification but has been obtained in electronic format from the US
 CC patent office at the following web site
 CC www.seqdata.uspto.gov/sequence.html; Document ID: 20040010132.
 XX
 XX Sequence 56 AA;
 QY 1 VXXXAEF 7
 Db 36 VAARAEF 42
 RESULT 38
 AAY79490
 ID AAY79490 standard; protein; 69 AA.
 XX AC AAY79490;
 XX 01-AUG-2000 (first entry)
 DT Human testis-specific 22P4F11 partial polypeptide.
 DE Human testis-specific 22P4F11 partial polypeptide.
 XX 22P4F11; human; testis; prostate cancer; diagnosis; therapy; marker;
 KW vaccine.
 KW Homo sapiens.
 OS WO200018925-A1.
 PN 06-APR-2000.
 XX 30-SEP-1999; 99WO-US023005.
 PF 30-SEP-1998; 98US-0102572P.
 PR 28-JUL-1999; 99US-0146584P.
 XX (UROC-) UROGENESYS INC.
 PA (AFAR/) AFAR D E.
 PA (HUBE/) HUBERT R S.
 PA (MITC/) MITCHELL S C.
 XX Afar DE, Hubert RS, Mitchell SC;
 PI WPI; 2000-303452/36.
 DR N-PSDB; AA294894.
 XX Novel testes-specific gene 22P4F11 which is expressed in human prostate
 PT cancer and is useful as a diagnostic marker and/or therapeutic target for
 PT prostate cancer.
 XX

PS Example 1; Fig 1B; 54pp; English.

CC The present sequence is that of a human 22P4F11 polypeptides, as deduced
 CC from an isolated partial cDNA (see AA294894). The full-length 22P4F11
 CC sequence is provided in MAY79489. 22P4F11 is a testis-specific gene in
 CC normal tissues, and is also expressed in human prostate tumours, in some
 CC cases at high levels. The 22P4F11 transcript and/or protein may represent
 CC a useful diagnostic marker and/or therapeutic target for prostate cancer.
 CC Methods of using 22P4F11 polynucleotides, polypeptides and antibodies for
 CC the diagnosis and treatment of cancers expressing 22P4F11, especially
 CC prostate cancer, are provided, as well as vaccines that prevent
 CC development of such cancers

XX SQ Sequence 69 AA;

Query Match 81.8%; Score 18; DB 3; Length 69;
 Best Local Similarity 42.9%; Pred. No. 8.1e+02;
 Matches 3; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
 : ||||
 Db 1 ISTTAEF 7

RESULT 39

ABP06455

ID ABP06455 standard; protein; 73 AA.

XX AC ABP06455;

XX 24-JUN-2002 (first entry)

DE Human ORFX protein sequence SEQ ID NO:12892.

XX Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis;
 XX hyperproliferative disorder; psoriasis; benign tumour; haemorrhage;
 KW degenerative disorder; osteoarthritis; neurodegenerative disorder;
 KW cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;
 KW hypertension; hypothyroidism; cholesterol ester storage disease;
 KW immune deficiency; immune disorder; infectious disease;
 KW autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;
 KW myasthenia gravis.

XX Homo sapiens.

XX WO200192523-A2.

XX 06-DEC-2001.

XX 29-MAY-2001; 2001WO-US010836.

XX 30-MAY-2000; 2000US-0206132P.

XX 29-AUG-2000; 2000US-0228716P.

XX (CURA-) CURAGEN CORP.

XX Shimkets RA, Leach MD;
 PI WPI; 2002-106308/14.
 XX N-PSDB; ABN22207.

XX Novel human polypeptides and polynucleotides useful for diagnosing,
 PT preventing and treating cardiovascular disease, neurodegenerative,
 PT hyperproliferative disorders and autoimmune disorders.

XX Disclosure; SEQ ID NO 12892; 1037pp; English.

CC The present invention describes substantially purified human proteins
 CC (referred to as open reading frame, ORFX, where X is 1-11491 (see Table 1
 CC in the specification). ABN15762 to ABN27252 encode the human ORFX
 CC proteins given in ABP00010 to ABP11500. ORFX proteins are useful for
 CC treating or preventing a pathology associated with an ORFX-associated
 CC disorder in humans, and in the manufacture of a medicament for treating a

CC syndrome associated with ORFX-associated disorder. ORFX polynucleotide
 CC sequences can be used in gene therapy. ORFX sequences can be used in the
 CC treatment of cancer, hyperproliferative disorders, cirrhosis of liver,
 CC psoriasis, benign tumours, keloid, degenerative disorders, haemorrhage,
 CC osteoarthritis, neurodegenerative disorders, disorders related to organ
 CC transplantation, cardiovascular diseases, diabetes mellitus, systemic
 CC lupus erythematosus, hypertension, hypothyroidism, cholesterol ester
 CC storage disease, various immune deficiencies and disorders, infectious
 CC diseases, autoimmune disorders such as multiple sclerosis, rheumatoid
 CC arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host
 CC disease and autoimmune inflammatory eye disease. ORFX proteins are also
 CC useful for treating burns, incisions, ulcers, for treating osteoporosis,
 CC bone degenerative disorders, or periodontal disease, and for gut
 CC protection or regeneration and treatment of lung or liver fibrosis,
 CC reperfusion injury in various tissues and conditions resulting from
 CC systemic cytokine damage. N.B. The sequence data for this patent did not
 CC form part of the printed specification, but was obtained in electronic
 CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 73 AA;

Query Match 81.8%; Score 18; DB 5; Length 73;
 Best Local Similarity 57.1%; Pred. No. 8.6e+02;
 Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
 : ||||
 Db 7 VATVAEF 13

RESULT 40

ADB09905

ID ADB09905 standard; protein; 83 AA.

XX AC ADB09905;

XX 20-NOV-2003 (first entry)

DE Alloiooccus otitis antigenic protein SEQ ID NO:3718.

XX Alloiooccus otitis; antigenic protein; immunogenic; immunisation;
 KW gene therapy; Gram-positive bacterium; infection.

XX Alloiooccus otitis.

XX WO2003048304-A2.

XX 12-JUN-2003.

XX 25-NOV-2002; 2002WO-US036123.

XX 29-NOV-2001; 2001US-0333777P.

XX 18-NOV-2002; 2002US-0426742P.

XX (AMHP) WYETH HOLDINGS CORP.

XX Fletcher LD, McMichael JC, Russell DP, Zagursky RJ;
 PI WPI; 2003-505284/47.
 XX N-PSDB; ADB09908.

XX New Alloiooccus otitis polynucleotides and polypeptides, useful for
 PT treating and diagnosing diseases, drug screening assays and monitoring of
 PT effects during drug clinical trials.

XX Claim 33; SEQ ID NO 3718; 1019pp; English.

CC The present invention describes an isolated polynucleotide (I) of
 CC Alloiooccus otitis genomic DNA, which encodes an antigenic protein.
 CC Alloiooccus otitis is a Gram-positive bacterium. Also described: (1)
 CC an isolated polypeptide that is encoded by the polynucleotide (I); (2) an
 CC expression vector comprising the novel isolated polynucleotide (I), its
 CC complement, degenerate variant or fragment; (3) a genetically engineered

CC host cell, transfected, transformed or infected with the vector of (2);
 CC (4) an antibody specific for the polypeptide of (1); (5) an immunogenic
 CC composition comprising the polypeptide, its complement, biological
 CC equivalent or fragment, or the polynucleotide that is comprised in the
 CC expression vector; (6) a pharmaceutical composition comprising the
 CC polypeptide of (1) and a carrier; (7) a protein chip comprising an array
 CC of the polypeptides of (1), their biological equivalent or fragment; (8)
 CC immunising against *Alloccoccus* otitis by administering to a host the
 CC immunogenic composition; (9) detecting and/or identifying *Alloccoccus*
 CC otitis in the biological sample; (10) a kit comprising a container
 CC containing the novel polynucleotide, its degenerate variant or fragment,
 CC or the antibody of (4); and (11) producing a polypeptide by culturing the
 CC genetically engineered host cell under conditions suitable to produce the
 CC polypeptide from the culture. (1) can be used in gene therapy. The
 CC polynucleotides, polypeptides, antibodies and compositions of the present
 CC invention can be used for treating and diagnosing diseases, drug
 CC screening assays and monitoring of effects during drug clinical trials.
 CC The polynucleotides are useful for expressing and detecting *Alloccoccus*
 CC otitis. The present sequence represents an *Alloccoccus* otitis
 CC antigen protein from the present invention.

XX
 SQ Sequence 83 AA;

Query Match 81.8%; Score 18; DB 6; Length 83;
 Best Local Similarity 57.1%; Pred. No. 9.9e+02;
 Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
 DB 21 VALAAEF 27

RESULT 41
 ADS21033
 ID ADS21033 standard; protein; 93 AA.

AC ADS21033;
 XX
 DT 02-DEC-2004 (first entry)
 XX
 DE Bacterial polypeptide #10066.

XX Recombinant DNA construct; transformed plant; improved plant property;
 KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
 KW pathogen tolerance; pest tolerance; plant disease resistance;
 KW cell cycle pathway modification; plant growth regulator;
 KW homologous recombination; seed oil yield; protein yield; carbohydrate;
 KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
 KW bacterial polypeptide.

XX Bacteria.

XX US2003233675-A1.

XX 18-DEC-2003.

XX 20-FEB-2003; 2003US-00369493.

XX 21-FEB-2002; 2002US-0360039P.

XX (CAOY/) CAO Y.
 PA (HINK/) HINKLE G J.
 PA (SLAT/) SLATER S C.
 PA (CHEN/) CHEN X.
 PA (GOLD/) GOLDMAN B S.

XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
 PI WPI; 2004-061375/06.

XX
 XX
 XX
 XX
 XX

PT New recombinant DNA construct comprising a promoter positioned to provide
 PT for expression of a polynucleotide encoding a polypeptide from a
 PT microbial source, useful for producing plants with improved properties.

XX Claim 1; SEQ ID NO 10066; 122pp; English.
 PS
 XX
 CC The invention relates to a recombinant DNA construct comprising a
 CC promoter functional in a plant cell, where the promoter is positioned to
 CC provide for expression of a polynucleotide encoding a polypeptide from a
 CC microbial source. The invention also relates to a transformed plant
 CC comprising the recombinant DNA construct and a method of producing a
 CC transformed plant having an improved property. The plant is a crop plant
 CC such as maize or soybean. The method of producing a transformed plant
 CC having an improved property comprises transforming a plant with the
 CC recombinant DNA construct and growing the transformed plant, where the
 CC polynucleotide or polypeptide is useful for improving plant properties.
 CC The recombinant DNA construct is useful for producing plants with
 CC improved plant properties, e.g. improved cold, heat or drought tolerance,
 CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
 CC increased resistance to plant disease, better growth rate by modification
 CC of the cell cycle pathway with plant growth regulators, increased rate of
 CC homologous recombination, modified seed oil or protein yield and/or
 CC content, improved yield by modification of carbohydrate, nitrogen or
 CC phosphorus use and/or uptake, by modification of photosynthesis or by
 CC providing improved plant growth and development under at least one stress
 CC condition, improved lignin production or improved galactomannan
 CC production. This sequence represents a bacterial polypeptide used in the
 CC scope of the invention. Note: The sequence data for this patent did not
 CC form part of the printed specification but was obtained in electronic
 CC format from USPTO at seqdata.uspto.gov/sequence.html.

XX Sequence 93 AA;

Query Match 81.8%; Score 18; DB 8; Length 93;
 Best Local Similarity 57.1%; Pred. No. 1.1e+03;
 Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
 DB 22 VRTAAEF 28

RESULT 42
 ABP05524
 ID ABP05524 standard; protein; 112 AA.

XX AC ABP05524;

XX 25-JUN-2002 (first entry)

XX Human ORFX protein sequence SEQ ID NO:11030.

XX Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis;
 KW hyperproliferative disorder; psoriasis; benign tumour; haemorrhage;
 KW degenerative disorder; osteoarthritis; neurodegenerative disorder;
 KW cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;
 KW hypertension; hypothyroidism; cholesterol ester storage disease;
 KW immune deficiency; immune disorder; infectious disease;
 KW autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;
 KW myasthenia gravis.

XX Homo sapiens.

XX WO200192523-A2.

XX 06-DEC-2001.

XX 29-MAY-2001; 2001WO-US010836.

XX 30-MAY-2000; 2000US-0206132P.

XX 29-AUG-2000; 2000US-0228716P.

XX (CURA-) CURAGEN CORP.

XX Shimkets RA, Leach MD;

XX

```

DR WPI; 2002-106308/14.
DR N-PSDB; ABN21276.
XX
PT Novel human polypeptides and polynucleotides useful for diagnosing,
PT preventing and treating cardiovascular disease, neurodegenerative,
PT hyperproliferative disorders and autoimmune disorders.
XX
PS Disclosure; SEQ ID NO 11030; 1037pp; English.
XX
CC The present invention describes substantially purified human proteins
CC (referred to as open reading frame, ORFX, where X is 1-11491 (see Table 1
CC in the specification) ABN15762 to ABN27252 encode the human ORFX
CC proteins given in ABP00010 to ABP1500. ORFX proteins are useful for
CC treating or preventing a pathology associated with an ORFX-associated
CC disorder in humans, and in the manufacture of a medicament for treating a
CC syndrome associated with ORFX-associated disorder. ORFX polynucleotide
CC sequences can be used in gene therapy. ORFX sequences can be used in the
CC treatment of cancer, hyperproliferative disorders, cirrhosis of liver,
CC psoriasis, benign tumours, keloid, degenerative disorders, haemorrhage,
CC osteoarthritis, neurodegenerative disorders, disorders related to organ
CC transplantation, cardiovascular diseases, diabetes mellitus, systemic
CC lupus erythematosus, hypertension, hypothyroidism, cholesterol ester
CC storage disease, various immune deficiencies and disorders, infectious
CC diseases, autoimmune disorders such as multiple sclerosis, rheumatoid
CC arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host
CC disease and autoimmune inflammatory eye disease. ORFX proteins are also
CC useful for treating burns, incisions, ulcers, for treating osteoporosis,
CC bone degenerative disorders, or periodontal disease, and for gut
CC protection or regeneration and treatment of lung or liver fibrosis,
CC reperfusion injury in various tissues and conditions resulting from
CC systemic cytokine damage. N.B. the sequence data for this patent did not
CC form part of the printed specification, but was obtained in electronic
CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 112 AA;

Query Match      81.8%; Score 18; DB 5; Length 112;
Best Local Similarity 42.9%; Pred. No. 1.4e+03;
Matches 3; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
Db 59 IATTAEF 65

RESULT 43
AAB29900
ID AAB29900 standard; protein; 121 AA.
AC AAB29900;
XX
DT 09-FEB-2001 (first entry)
DE Human secreted protein BLAST search protein SEQ ID NO: 158.
XX
KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
KW anti-allergic; hepatotropic; antidiabetic; anti-inflammatory; anti-ulcer;
KW vulnerrary; anticonvulsant; antibacterial; antifungal; antiparasitic;
KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
KW neurological disease; infection; human; secreted protein.
XX
OS Homo sapiens.
XX
PN WO200061779-A1.
XX
PD 19-OCT-2000.
XX
PF 06-APR-2000; 2000WO-US009068.
XX
PR 09-APR-1999; 99US-0128699P.
XX
PR 20-JAN-2000; 2000US-0177050P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Ruben SM, Komatsoulis G;
WPI; 2000-647424/62.
Isolated nucleic acid molecule encoding a human secreted protein is used
in preventing, treating or ameliorating a medical condition.
Disclosure; Page 479; 495pp; English.
The invention relates to the isolation of genes AAC63410-C63458 encoding
the human secreted proteins AAB29802-B29850. This sequence represents a
fragment of the protein encoded by the gene given in the descriptor line.
The sequence is used as a query sequence for doing BLASTX searches to
determine homologous sequence to the protein. The genes and proteins are
useful for preventing, ameliorating or treating medical conditions, e.g.
by protein or gene therapy. The genes are isolated from a range of human
tissues disclosed in the specification. The nucleic acids, proteins,
antibodies and (ant)agonists are useful in the diagnosis, treatment and
prevention of: (a) cancer, e.g. breast and ovarian cancer, and other
cancers of the adrenal gland, bone, bone marrow, breast, gastrointestinal
tract, liver, lung, or urogenital; (b) immune disorders e.g. Addison's
disease, allergies, autoimmune haemolytic anaemia, autoimmune
thyroiditis, diabetes mellitus, Crohn's disease, multiple sclerosis,
rheumatoid arthritis and ulcerative colitis; (c) cardiovascular disorders
such as myocardial ischaemias; (d) wound healing; (e) neurological
diseases e.g. cerebral anoxia and epilepsy; and (f) infectious diseases
such as viral, bacterial, fungal and parasitic infections
XX
SQ Sequence 121 AA;

Query Match      81.8%; Score 18; DB 3; Length 121;
Best Local Similarity 42.9%; Pred. No. 1.5e+03;
Matches 3; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
Db 49 ISTTAEF 55

RESULT 44
ADF76715
ID ADF76715 standard; protein; 137 AA.
XX
AC ADF76715;
XX
DT 26-FEB-2004 (first entry)
DE Novel human secreted and transmembrane protein SeqID 390.
XX
KW human; PRO; membrane bound protein; membrane bound receptor;
KW cell proliferation; cell migration; cell differentiation;
KW mitogenic factor; survival factor; cytotoxic factor;
KW differentiation factor; neuropeptide; hormone; cell receptor;
KW receptor-ligand interaction; cytostatic; chondrocyte; tumour.
XX
OS Homo sapiens.
XX
PN WO2003072035-A2.
XX
PD 04-SEP-2003.
XX
PF 21-FEB-2003; 2003WO-US005241.
XX
PR 22-FEB-2002; 2002US-0359461P.
XX
PA (GETH ) GENENTECH INC.
XX
PI Bodary SC, Clark H, Hunte B, Jackman JK, Schoenfeld JR;
PI Williams PM, Wood WI, Wu TD;
XX
XX WPI; 2003-721702/68.
DR N-PSDB; ADF76714.

```

XX New PRO polypeptides, useful for diagnosing and treating an immune
PT related disorder, e.g. systemic lupus erythematosus, rheumatoid
PT arthritis, osteoarthritis, juvenile chronic arthritis, thyroiditis or
PT diabetes mellitus.
XX
XX Claim 10; SEQ ID NO 390; 918pp; English.
XX
XX This invention relates to novel nucleic acids encoding human PRO secreted
CC and transmembrane proteins. Extracellular proteins play important roles
CC in the formation, differentiation and maintenance of multicellular
CC organisms. The fate of many individual cells (for example proliferation,
CC migration or differentiation) is typically governed by information
CC received from other cells and the immediate environment. The information
CC is often transmitted by secreted polypeptides (for example mitogenic
CC factors, survival factors, cytotoxic factors, differentiation factors,
CC neuropeptides and hormones) which are received and interpreted by diverse
CC cell receptors or membrane bound proteins. These membrane bound proteins
CC and receptors may be of use as pharmaceutical and diagnostic agents, such
CC as in the blocking of receptor-ligand interactions. The current invention
CC provides the amino acid sequences of novel human membrane bound receptors
CC and proteins, along with the cDNA sequences encoding them. The novel
CC proteins of the invention may have cytosolic activities through the
CC stimulation of chondrocytes. The nucleic acids of the invention may be
CC useful for the manufacture of a medicament for diagnosing or treating a
CC tumour in a mammal. In addition, they may be useful for measuring or
CC detecting the expression of a tumour associated gene. The present
CC sequence is the amino acid sequence of a human PRO protein of the
CC invention.
XX
XX SQ Sequence 137 AA;
Query Match 81.8%; Score 18; DB 7; Length 137;
Best Local Similarity 57.1%; Pred. No. 1.7e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 VXXXXAEF 7
Db 7 VITAEF 13

RESULT 45
AA99815
ID AA99815 standard; protein; 139 AA.
XX
XX AA99815;
XX
XX 19-SEP-2000 (first entry)
XX
XX Soybean partial protein phosphatase 2A regulatory subunit B.
XX
XX Soybean; protein phosphatase 2A; protein phosphorylation modulation;
KW transgenic plant; gene therapy.
XX
XX Glycine max.
XX
XX Key Location/Qualifiers
FH Misc-difference 138
FT /label= unknown
FT /note= "encoded by NAN"
XX
XX WO200036121-A2.
XX
XX 22-JUN-2000.
XX
XX 15-DEC-1999; 99WO-US029823.
XX
XX 16-DEC-1998; 98US-0112541P.
XX
XX (DUPO) DU PONT DE NEMOURS & CO E I.
XX
XX Famodu OO, Miao G, Sakai H, Lee J, Rafalski JA, Klein TM;
XX

DR WPI; 2000-431599/37.
DR N-PSDB; AAA48570.
XX
XX Polynucleotides encoding plant protein phosphatase useful for modulating
PT reversible protein phosphorylation in plants.
XX
XX Claim 12; Page 48-49; 73pp; English.
XX
XX The present sequence is soybean protein phosphatase 2A regulatory subunit
CC B. The nucleotide sequence encoding this protein was identified in clone
CC src2c.pk023.19 of a cDNA library made from soybean eight-day-old root
CC infected with cyst nematode. BLAST analysis showed that the sequence
CC encoded protein phosphatase 2A regulatory subunit B. The nucleotide
CC sequence may be used for the recombinant production of the protein in
CC vivo, e.g. via a gene therapy protocol, or in vitro, e.g. in fermentation
CC culture. The protein may then be used to modulate the process of
CC reversible protein phosphorylation in plants. It may be used directly to
CC supplement a plant's own production of the enzyme or to rectify mutations
CC that result in the expression of inactive protein. The protein may also
CC be used to test for modulators of protein phosphorylation which may be
CC used to alter the activity of the enzyme
XX
XX SQ Sequence 139 AA;
Query Match 81.8%; Score 18; DB 3; Length 139;
Best Local Similarity 57.1%; Pred. No. 1.8e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 VXXXXAEF 7
Db 2 VITSAEF 8

RESULT 46
AAU31774
ID AAU31774 standard; protein; 144 AA.
XX
XX AAU31774;
XX
XX 18-DEC-2001 (first entry)
XX
XX Novel human secreted protein #2265.
DE
XX Human; vaccination; gene therapy; nutritional supplement;
XX stem cell proliferation; haematopoiesis; nerve tissue regeneration;
KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.
XX
XX Homo sapiens.
XX
XX WO200179449-A2.
XX
XX 25-OCT-2001.
PD
XX 16-APR-2001; 2001WO-US008656.
XX
XX 18-APR-2000; 2000US-00552929.
XX
XX 26-JAN-2001; 2001US-00770160.
XX
XX (HYSE-) HYSEQ INC.
PA
XX Tang YT, Liu C, Drmanac RT;
XX
XX WPI; 2001-611725/70.
XX
XX Nucleic acids encoding a range of human polypeptides, useful in genetic
PT vaccination, testing and therapy.
XX
XX Claim 20; Page 507; 765pp; English.
XX
XX The invention relates to novel human secreted polypeptides. The
CC polypeptides and antibodies to the polypeptides are useful for
CC determining the presence of or predisposition to a disease associated
CC with altered levels of polypeptide. The polypeptides are also useful for

CC identifying agents (agonists and antagonists) that bind to them. Cells
 CC expressing the proteins are useful for identifying a therapeutic agent
 CC for use in treatment of a pathology related to aberrant expression or
 CC physiological interactions of the polypeptide. Vectors comprising the
 CC nucleic acids encoding the polypeptides and cells genetically engineered
 CC to express them are also useful for producing the proteins. The proteins
 CC are useful in genetic vaccination, testing and therapy, and can be used
 CC as nutritional supplements. They may be used to increase stem cell
 CC proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon
 CC and/or nerve tissue growth or regeneration; immune suppression and/or
 CC stimulation; as anti-inflammatory agents; and in treatment of leukaemias.
 CC AAU29510-AAU3304 represent the amino acid sequences of novel human
 CC secreted proteins of the invention
 XX
 SQ Sequence 144 AA;

Query Match 81.8%; Score 18; DB 4; Length 144;
 Best Local Similarity 57.1%; Pred. No. 1.8e+03;
 Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXXAEP 7
 | |||
 Db 115 VITAAEF 121

RESULT 47
 ABB55379
 ID ABB55379 standard; protein; 148 AA.

XX
 AC ABB55379;

XX 29-AUG-2003 (revised)

DT 16-MAY-2002 (first entry)

XX Lactococcus lactis protein yvaC.

XX Biosynthesis; biodegradation; lactic bacterium; yogurt; cheese.

XX Lactococcus lactis; IL1403.

XX FR280746-A1.

PN 12-OCT-2001.

XX 11-APR-2000; 2000FR-00004630.

XX 11-APR-2000; 2000FR-00004630.

XX (INRG) INRA INST NAT RECH AGRONOMIQUE.

XX Bolotine A, Sorokine A, Renault P, Ehrlich SD;

XX WPI; 2002-043418/06.

XX New nucleotide sequence useful in the identification or Lactococcus
 PT lactis and related species.

XX Claim 6; SEQ ID NO 2081; 2504pp; French.

XX The present invention is related to a Lactococcus lactis nucleotide
 CC sequence (ABA90521) and related proteins (ABB53300-ABB55621). The nucleic
 CC acid sequence is useful in the detection and/or amplification of nucleic
 CC acid sequence, particularly to identify Lactococcus lactis or related
 CC species. The proteins of the invention are useful for the biosynthesis or
 CC biodegradation of a composition of interest. The invention helps research
 CC in lactic bacteria, particularly useful in the production of yogurt and
 CC cheese. Note: The sequence data for this patent is based on equivalent
 CC patent WO2001/77334 (published 18-OCT-2001) which is available in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences. (Updated on 29-AUG-2003 to
 CC standardise OS field)

XX Sequence 148 AA;

Query Match 81.8%; Score 18; DB 5; Length 148;
 Best Local Similarity 57.1%; Pred. No. 1.9e+03;
 Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXXAEP 7
 | |||
 Db 69 VSTLAEP 75

RESULT 48

AA334523

ID AAB34523 standard; protein; 155 AA.

XX
 AC AAB34523;

XX 02-FEB-2001 (first entry)

XX Human secreted protein BLAST search protein SEQ ID NO: 141.

XX Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
 KW anti-allergic; hepatotropic; antidiabetic; anti-inflammatory; antitumor;
 KW vulnery; anticonvulsant; antibacterial; antifungal; antiparasitic;
 KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; human; secreted protein.

XX Homo sapiens.

XX WO200056767-A1.

PN 28-SEP-2000.

XX 16-MAR-2000; 2000WO-US0066828.

XX 19-MAR-1999; 99US-0125358P.

XX 08-DEC-1999; 99US-0169616P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Ruben SM, Komatsoulis G;

XX WPI; 2000-602216/57.

XX Nucleic acid molecules encoding human secreted proteins, used in
 PT preventing, treating or ameliorating a disorder, e.g. Alzheimer's and
 PT Parkinson's diseases and cancers.

XX Disclosure; Page 373; 384pp; English.

XX The invention relates to the isolation of genes AAC59624-C59669 encoding
 CC the human secreted proteins AAB34439-B34484. This sequence represents a
 CC peptide fragment homologous to the protein encoded by the gene isolated
 CC in the present invention. The sequence is a search result from a BLASTX
 CC homology search. The genes and proteins are useful for preventing,
 CC ameliorating or treating medical conditions, e.g. by protein or gene
 CC therapy. The genes are isolated from a range of human tissues disclosed
 CC in the specification. The nucleic acids, proteins, antibodies and
 CC (ant)agonists are useful in the diagnosis, treatment and prevention of:
 CC (a) cancer, e.g. breast and ovarian cancer, and other cancers of the
 CC adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver,
 CC lung, or urogenital; (b) immune disorders e.g. Addison's disease,
 CC allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,
 CC diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid
 CC arthritis and ulcerative colitis; (c) cardiovascular disorders such as
 CC myocardial ischaemia; (d) wound healing; (e) neurological diseases e.g.
 CC cerebral anoxia and epilepsy; and (f) infectious diseases such as viral,
 CC bacterial, fungal and parasitic infections

XX Sequence 155 AA;

Query Match 81.8%; Score 18; DB 3; Length 155;
 Best Local Similarity 57.1%; Pred. No. 2e+03;
 Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
|
44 VATAEF 50

Db

RESULT 49
ADS27899
ID ADS27899 standard; protein; 159 AA.
XX
AC ADS27899;
XX
DT 02-DEC-2004 (first entry)
XX
DE Bacterial polypeptide #16932.
XX
KW Recombinant DNA construct; transformed plant; improved plant property;
cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
pathogen tolerance; pest tolerance; plant disease resistance;
cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
bacterial polypeptide.
XX
OS Bacteria.
XX
PN US2003233675-A1.
XX
PD 18-DEC-2003.
XX
PF 20-FEB-2003; 2003US-00369493.
XX
PR 21-FEB-2002; 2002US-0360039P.
XX
PA (CAOY/) CAO Y.
(HINK/) HINKLE G J.
PA (SLAT/) SLATER S C.
PA (CHEN/) CHEN X.
PA (GOLD/) GOLDMAN B S.
XX
PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
WPI; 2004-061375/06.
XX
DR New recombinant DNA construct comprising a promoter positioned to provide
for expression of a polynucleotide encoding a polypeptide from a
microbial source, useful for producing plants with improved properties.
XX
PS Claim 1; SEQ ID NO 16932; 122pp; English.
XX
CC The invention relates to a recombinant DNA construct comprising a
promoter functional in a plant cell, where the promoter is positioned to
provide for expression of a polynucleotide encoding a polypeptide from a
microbial source. The invention also relates to a transformed plant
comprising the recombinant DNA construct and a method of producing a
transformed plant having an improved property. The plant is a crop plant
such as maize or soybean. The method of producing a transformed plant
having an improved property comprises transforming a plant with the
recombinant DNA construct and growing the transformed plant, where the
polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
improved plant properties, e.g. improved cold, heat or drought tolerance,
tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
increased resistance to plant disease, better growth rate by modification
of the cell cycle pathway with plant growth regulators, increased rate of
homologous recombination, modified seed oil or protein yield and/or
content, improved yield by modification of carbohydrate, nitrogen or
phosphorus use and/or uptake, by modification of photosynthesis or by
providing improved plant growth and development under at least one stress
condition. Improved lignin production or improved galactomannan
production. This sequence represents a bacterial polypeptide used in the
scope of the invention. Note: The sequence data for this patent did not
form part of the printed specification but was obtained in electronic

CC format from USPTO at seqdata.uspto.gov/sequence.html.
XX
SQ Sequence 159 AA;
Query Match 81.8%; Score 18; DB 8; Length 159;
Best Local Similarity 57.1%; Pred. NO. 2e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 VXXAAEF 7
|
Db 142 VAAAEF 148

RESULT 50
AAU67631
ID AAU67631 standard; protein; 162 AA.
XX
AC AAU67631;
XX
DT 27-FEB-2002 (first entry)
XX
DE Propionibacterium acnes immunogenic protein #28527.
XX
KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
uvelitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant.
XX
OS Propionibacterium acnes.
XX
PN WO200181581-A2.
XX
PD 01-NOV-2001.
XX
PF 20-APR-2001; 2001WO-US012865.
XX
PR 21-APR-2000; 2000US-0199047P.
PR 02-JUN-2000; 2000US-0208841P.
PR 07-JUL-2000; 2000US-0216747P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX
DR WPI; 2001-616774/71.
DR N-PSDB; AAS59604.
XX
PT Propionibacterium acnes polypeptides and nucleic acids useful for
vaccinating against and diagnosing infections, especially useful for
treating acne vulgaris.
XX
PS Example 1; SEQ ID NO 28826; 1069pp; English.
XX
CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
polypeptides. The proteins and their associated DNA sequences are used in
the treatment, prevention and diagnosis of medical conditions caused by
P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
pustulosis, hypertosis and osteomyelitis), uvelitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
nervous system, however it is particularly involved in the inflammatory
lesions associated with acne vulgaris. A method for detecting the
presence or absence of P. acnes in a patient comprises contacting a
sample with a binding agent that binds to the proteins of the invention
and determining the amount of bound protein in the sample. The
polypeptides may be used as antigens in the production of antibodies
specific for P. acnes proteins. These antibodies can be used to
downregulate expression and activity of P. acnes polypeptides and
therefore treat P. acnes infections. The antibodies may also be used as
diagnostic agents for determining P. acnes presence, for example, by
enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
this patent did not form part of the printed specification, but was
obtained in electronic format directly from WIPO at

```

CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 162 AA;

Query Match      81.8%; Score 18; DB 4; Length 162;
Best Local Similarity 57.1%; Pred. No. 2.1e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXXAEF 7
Db 100 VAVTAEF 106

RESULT 51
ID ABM64150 standard; protein; 162 AA.
XX
AC ABM64150;
XX
DT 20-OCT-2003 (first entry)
XX
DE Propionibacterium acnes predicted ORF-encoded polypeptide #28926.
XX
DE Acne vulgaris; antiseborrheic; dermatological; antibacterial;
KW immunostimulant; immune response; vaccine.
XX
OS Propionibacterium acnes.
XX
PN WO2003033515-A1.
XX
PD 24-APR-2003.
XX
PF 11-OCT-2002; 2002WO-US032727.
XX
PR 15-OCT-2001; 2001US-00978825.
XX
PA (CORI-) CORIXA CORP.
XX
PI Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;
FI Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;
PI Barth B, Valliave-Douglas J;
XX
DR WPI; 2003-381789/36.
DR N-PSDB; ACF64533.
XX
XX
PT New Propionibacterium acnes polypeptides and polynucleotides encoding the
PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
PT or for stimulating an immune response specific for a P. acnes protein.
PS Example 1; SEQ ID NO 28926; 1481pp; English.
XX

The invention relates to an isolated polynucleotide (ACF64435-ACF64733)
CC encoding a Propionibacterium acnes protein. The invention also relates to
CC polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to
CC immunogenic fragments of P. acnes polypeptides. The invention
CC additionally encompasses expression vectors and host cells comprising a
CC polynucleotide of the invention; antibodies against polypeptides of the
CC invention; fusion proteins comprising a polypeptide of the invention; a
CC method for stimulating an immune response specific for a P. acnes
CC polypeptide and an isolated T cell population comprising T cells prepared
CC via this method; a vaccine composition (comprising P. acnes polypeptides,
CC polynucleotides, antibodies, fusion proteins), T cell populations, or
CC antigen-presenting cells that express the polypeptide; a method and kit
CC for detecting or determining the presence or absence of P. acnes in a
CC patient; and a method for inhibiting the development of P. acnes in a
CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion
CC proteins, T cell populations or antigen-presenting cells that express the
CC polypeptides are useful for diagnosing, preventing or treating acne
CC vulgaris, or for stimulating an immune response specific for a P. acnes
CC protein. The polynucleotides can also be used as probes or primers for
CC nucleic acid hybridisation. The vaccine composition is useful for the
CC stimulation of an immune response against P. acnes, or for treating acne,
CC and the kit is useful for performing a diagnostic assay. The present

```

```

CC sequence represents a polypeptide predicted to be encoded by an ORF (open
CC reading frame) contained within the P. acnes polynucleotides of the
CC invention. Note: the sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 162 AA;

Query Match      81.8%; Score 18; DB 6; Length 162;
Best Local Similarity 57.1%; Pred. No. 2.1e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXXAEF 7
Db 100 VAVTAEF 106

RESULT 52
ID AAB28213 standard; protein; 169 AA.
XX
AC AAB28213;
XX
DT 30-JAN-2001 (first entry)
XX
DE Novel human protein #11.
XX
KW Cytostatic; vaccine; human; breast tumour; antigen; breast cancer.
XX
OS Homo sapiens.
XX
PN WO200052165-A2.
XX
PD 08-SEP-2000.
XX
PF 29-FEB-2000; 2000WO-US005431.
XX
PR 04-MAR-1999; 99US-00262505.
PR 19-MAR-1999; 99US-00272886.
PR 17-SEP-1999; 99US-00396313.
XX
PA (CORI-) CORIXA CORP.
XX
PI Lodes MJ;
XX
WPI; 2000-572184/53.
DR N-PSDB; AAC69697.
XX
PT Breast tumor antigen polypeptides and polynucleotides, useful for
PT manufacturing vaccines and compositions for treating, diagnosing, and
PT monitoring breast cancer.
XX
XX
XX Example; Fig 2; 140pp; English.
XX
The present invention relates to immunogenic portions of new human breast
CC tumour antigens (AAB28183-B28214) and their coding sequences (AAC69645-
CC C69804). The breast tumour antigen polypeptides of the present invention
CC and their coding sequences are useful for inhibiting the development of
CC breast cancer in a patient. The breast tumour antigen polypeptides and
CC polynucleotides may be used in vaccines and pharmaceutical compositions
CC for treating breast cancer, and for diagnosing and monitoring the cancer.
CC The present sequence is a immunogenic portion for one such human breast
CC cancer tumour antigen
XX
XX
SQ Sequence 169 AA;

Query Match      81.8%; Score 18; DB 3; Length 169;
Best Local Similarity 57.1%; Pred. No. 2.2e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXXAEF 7
Db 51 VYSSAEF 57

```



```
RESULT 53
AD167204
ID AD167204 standard; protein; 179 AA.
XX
XX AC AD167204;
XX
XX DT 22-APR-2004 (first entry)
XX
XX DE Lactobacillus rhamnosus polypeptide sequence #78.
XX
XX KW Lactic acid bacteria; vaccine; bacterial infection; microbe;
XX milk-derived product; food product; food additive;
XX KW nutritional supplement; bioactive substance; probiotic supplement;
XX flavour; aroma; texture; nutritional value; food; microorganism;
XX antibacterial.
XX
XX OS Lactobacillus rhamnosus; strain HN001.
XX
XX PN US2004009490-A1.
XX
XX PD 15-JAN-2004.
XX
XX PF 03-OCT-2002; 2002US-00264213.
XX
XX PR 09-AUG-1999; 99US-0147852P.
XX PR 09-AUG-1999; 99US-0147853P.
XX PR 01-SEP-1999; 99US-0152031P.
XX PR 01-SEP-1999; 99US-0152032P.
XX PR 08-AUG-2000; 2000US-00634238.
XX PR 02-OCT-2001; 2001US-00971536.
XX
XX PA (GENE-) GENESIS RES & DEV CORP LTD.
XX
XX PI Glenn M, Havukkala IU, Lubbers M, Dekker J;
XX
XX WIPI; 2004-090459/09.
XX
XX DR N-PSDB; AD167079.
XX
XX PT New polynucleotide from Lactobacillus rhamnosus HN001 strain, useful for
XX preparing a vaccine against bacterial infections or for modifying the
XX flavor, aroma or nutritional benefits of a bioactive or probiotic
XX supplement product.
XX
XX PS Claim 21; SEQ ID NO 199; 54pp; English.
XX
XX CC The present invention relates to the isolation of novel lactic acid
XX bacteria (Lactobacillus rhamnosus strain HN001) polynucleotide sequences,
XX and the proteins encoded by them. Also disclosed are oligonucleotide
XX probes and primers, and genetic constructs comprising the polynucleotide
XX sequences of the invention. The polynucleotide sequences are useful for
XX preparing a vaccine against bacterial infections or for improving the
XX properties of microbes used in the manufacture of milk-derived products,
XX food products, food additives, nutritional supplements, bioactive
XX substances or probiotic supplements, and for modifying the flavour,
XX aroma, texture and/or nutritional value of foods. They are also useful
XX for identifying microorganisms having a trait associated with the
XX polynucleotide. The present sequence represents a novel L. rhamnosus
XX polypeptide sequence of the invention. Note: The sequence data for this
XX patent did not form part of the printed specification. The complete
XX sequence data for this patent was obtained in electronic format directly
XX from the USPTO web site at seqdata.uspto.gov.
XX
XX SQ Sequence 179 AA;
XX
XX Query Match 81.8%; Score 18; DB 8; Length 179;
XX Best Local Similarity 57.1%; Pred. No. 2.3e+03;
XX Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
XX
XX QY 1 VXXXAEF 7
XX | |||
XX Db 162 VSTIAEF 168

RESULT 54
ABP29671
ID ABP29671 standard; protein; 193 AA.
XX
XX AC ABP29671;
XX
XX DT 02-JUL-2002 (first entry)
XX
XX DE Streptococcus polypeptide SEQ ID NO 8518.
XX
XX KW Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
XX group A streptococcus; Streptococcus pyogenes; antibacterial;
XX antiinflammatory; infection; vaccine; meningitis; gene therapy.
XX
XX OS Streptococcus agalactiae.
XX
XX PN WO200234771-A2.
XX
XX PD 02-MAY-2002.
XX
XX PF 29-OCT-2001; 2001WO-GB004789.
XX
XX PR 27-OCT-2000; 2000GB-00026333.
XX PR 24-NOV-2000; 2000GB-00028727.
XX PR 07-MAR-2001; 2001GB-00005640.
XX
XX PA (CHIR-) CHIRON SPA.
XX
XX PA (GENO-) INST GENOMIC RES.
XX
XX PI Telford J, Masighani V, Margarit Y RosI, Grandi G, Fraser C;
XX Tettelin H;
XX
XX WIPI; 2002-352536/38.
XX
XX DR N-PSDB; ABN70302.
XX
XX PT New Streptococcus protein for the treatment or prevention of infection or
XX disease caused by Streptococcus bacteria, such as meningitis, and for
XX detecting a compound that binds to the protein.
XX
XX PS Claim 1; Page 3962; 4525pp; English.
XX
XX CC The invention relates to a protein (ABP25413-ABP30895) from group B
XX streptococcus/GSS (Streptococcus agalactiae) or group A streptococcus/GAS
XX (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in
XX the specification. The proteins have antibacterial and antiinflammatory
XX activity. (I), nucleic acids encoding (I), ABN66044-ABN71526 and
XX antibodies that bind (I) are used in the manufacture of medicaments for
XX the treatment or prevention of infection or disease caused by
XX Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
XX Nucleic acids encoding (I) are used to detect Streptococcus in a
XX biological sample. (I) is used to determine whether a compound binds to
XX (I). A composition comprising (I) or a nucleic acid encoding (I), may be
XX used as a vaccine or diagnostic composition. The disease caused by
XX Streptococcus that is prevented or treated may be meningitis. Nucleic
XX acid encoding (I) may be used to recombinantly produce (I) and may be
XX used in gene therapy. Antibodies to (I) are used for affinity
XX chromatography, immunoassays, and distinguishing/identifying
XX Streptococcus proteins
XX
XX SQ Sequence 193 AA;
XX
XX Query Match 81.8%; Score 18; DB 5; Length 193;
XX Best Local Similarity 57.1%; Pred. No. 2.5e+03;
XX Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
XX
XX QY 1 VXXXAEF 7
XX | |||
XX Db 41 VTSIAEF 47

RESULT 55
```

ABP25742
ID ABP25742 standard; protein; 197 AA.
XX
AC ABP25742;
XX
DT 02-JUL-2002 (first entry)
XX
DE Streptococcus polypeptide SEQ ID NO 660.
XX
KW Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
KW group A streptococcus; Streptococcus pyogenes; antibacterial;
KW antiinflammatory; infection; vaccine; meningitis; gene therapy.
XX
OS Streptococcus agalactiae.
XX
PN WO200234771-A2.
XX
PD 02-MAY-2002.
XX
PF 29-OCT-2001; 2001WO-GB004789.
XX
PR 27-OCT-2000; 2000GB-00026333.
PR 24-NOV-2000; 2000GB-00028727.
PR 07-MAR-2001; 2001GB-00005640.
XX
XX (CHIR-) CHIRON SPA.
PA
PA (GENO-) INST GENOMIC RES.
XX
XX Telford J, Masignani V, Margazit Y, Rosi I, Grandi G, Fraser C;
PI Tettelin H;
XX
XX WPI; 2002-352536/38.
DR N-PSDB; ABN66373.
XX
XX New Streptococcus protein for the treatment or prevention of infection or
PT disease caused by Streptococcus bacteria, such as meningitis, and for
PT detecting a compound that binds to the protein.
XX
XX Claim 1; Page 3219; 4525pp; English.
XX
XX The invention relates to a protein (ABP25413-ABP30895) from group B
CC streptococcus/GAS (Streptococcus agalactiae) or group A streptococcus/GAS
CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in
CC the specification. The proteins have antibacterial and antiinflammatory
CC activity. (I), nucleic acids encoding (I), ABN66044-ABN71526 and
CC antibodies that bind (I) are used in the manufacture of medicaments for
CC the treatment or prevention of infection or disease caused by
CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
CC Nucleic acids encoding (I) are used to detect Streptococcus in a
CC biological sample. (I) is used to determine whether a compound binds to
CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be
CC used as a vaccine or diagnostic composition. The disease caused by
CC Streptococcus that is prevented or treated may be meningitis. Nucleic
CC acid encoding (I) may be used to recombinantly produce (I) and may be
CC used in gene therapy. Antibodies to (I) are used for affinity
CC chromatography, immunoassays, and distinguishing/identifying
CC Streptococcus proteins
XX
SQ Sequence 197 AA;
Query Match 81.8%; Score 18; DB 5; Length 197;
Best Local Similarity 57.1%; Pred. No. 2.6e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 1 VXXXXAEF 7
Db 45 VTSIAEF 51
RESULT 56
AAY34533
ID AAY34533 standard; protein; 204 AA.
XX

AC AAY34533;
XX
DT 27-AUG-2003 (revised)
DT 20-MAR-2003 (revised)
DT 25-AUG-1999 (first entry)
XX
DE Porphyromonas gingivalis protein PG64.
XX
KW Porphyromonas gingivalis; PG; periodontal disease; gingivitis; vaccine;
KW antigenic.
XX
OS Porphyromonas gingivalis.
XX
PN WO9929870-A1.
XX
PD 17-JUN-1999.
XX
PF 10-DEC-1998; 98WO-AU001023.
XX
PR 10-DEC-1997; 97AU-00000839.
PR 31-DEC-1997; 97AU-00001182.
PR 30-JAN-1998; 98AU-00001546.
PR 10-MAR-1998; 98AU-00002264.
PR 09-APR-1998; 98AU-00002911.
PR 23-APR-1998; 98AU-00003128.
PR 05-MAY-1998; 98AU-00003338.
PR 22-MAY-1998; 98AU-00003654.
PR 29-JUL-1998; 98AU-00004917.
PR 30-JUL-1998; 98AU-00004963.
PR 04-AUG-1998; 98AU-00005028.
XX
PA (CSLC-) CSL LTD.
XX
XX Ross BC, Barr IG, Patterson MA, Agius CT, Rothel LJ, Margetts MB;
PI Hocking DM, Webb EA;
XX
XX WPI; 1999-385613/32.
DR N-PSDB; AAX91751.
XX
XX Antigenic Porphyromonas gingivalis peptides for preventing gingivitis.
PT
XX Claim 1; Page 524; 588pp; English.
XX
XX AAX91536 to AAX91801 encode two hundred and sixty six antigenic
CC Porphyromonas gingivalis (PG) polypeptide sequences given in AAY34318 to
CC AAY34583. AAX91802 to AAX91989 represent PCR primers used in the
CC isolation of the PG polypeptides. The PG polypeptides have antibacterial
CC activity with a vaccine mechanism of action. The PG polypeptides can be
CC used as vaccines especially against Porphyromonas gingivalis. Probes can
CC be used to detect Porphyromonas gingivalis in standard hybridisation
CC assays. Porphyromonas gingivalis is involved in periodontal disease
CC especially gingivitis. (Updated on 20-MAR-2003 to correct PR field.)
XX (Updated on 27-AUG-2003 to correct OS field.)
SQ Sequence 204 AA;
Query Match 81.8%; Score 18; DB 2; Length 204;
Best Local Similarity 57.1%; Pred. No. 2.7e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 1 VXXXXAEF 7
Db 56 VGNAAEF 62
RESULT 57
ABO66356
ID ABO66356 standard; protein; 214 AA.
XX
AC ABO66356;
XX
DT 29-JUL-2004 (first entry)
XX

DE Klebsiella pneumoniae polypeptide seqid 12873.
 XX Recombinant expression vector; transcription regulatory element;
 KW Klebsiella pneumoniae protein; antibacterial; Vaccine.
 XX
 OS Klebsiella pneumoniae.
 PN US6610836-B1.
 XX
 XX 26-AUG-2003.
 XX
 PF 27-JAN-2000; 2000US-00489039.
 XX
 XX 29-JAN-1999; 99US-0117747P.
 XX
 XX (GENO-) GENOME THERAPEUTICS CORP.
 XX
 PI Breton GL, Osborne M;
 XX
 DR WPI; 2003-895346/82.
 DR N-PSDB; ACH99907.
 XX
 XX New nucleic acid encoding a Klebsiella pneumoniae polypeptide, useful for
 PT preparing a vaccine composition against Klebsiella pneumoniae.
 XX
 PS Disclosure; SEQ ID NO 12873; 932pp; English.
 XX
 CC The invention describes a new isolated nucleic acid encoding a Klebsiella
 CC pneumoniae polypeptide. Also described are: a recombinant expression
 CC vector comprising the nucleic acid, operably linked to a transcription
 CC regulatory element; and a cell comprising the recombinant expression
 CC vector. The nucleic acid is useful for preparing a vaccine composition
 CC against Klebsiella pneumoniae. This is the amino acid sequence of a
 CC Klebsiella pneumoniae polypeptide of the invention
 XX
 SQ Sequence 214 AA;
 Query Match 81.8%; Score 18; DB 7; Length 214;
 Best Local Similarity 57.1%; Pred. NO. 2.8e+03;
 Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 VXXXXAEF 7
 Db 186 VVSTAEF 192
 RESULT 58
 AAY34400
 ID AAY34400 standard; protein; 221 AA.
 XX
 AC AAY34400;
 XX
 XX 27-AUG-2003 (revised)
 DT 20-MAR-2003 (revised)
 DT 25-AUG-1999 (first entry)
 XX
 DE Porphyromonas gingivalis protein PG64.
 XX
 KW Porphyromonas gingivalis; PG; periodontal disease; gingivitis; vaccine;
 KW antigenic.
 XX
 OS Porphyromonas gingivalis.
 XX
 PN WO9929870-A1.
 XX
 XX 17-JUN-1999.
 XX
 PF 10-DEC-1998; 98WO-AU001023.
 XX
 XX 10-DEC-1997; 97AU-00000839.
 PR 31-DEC-1997; 97AU-00001182.
 PR 30-JAN-1998; 98AU-00001546.
 PR 10-MAR-1998; 98AU-00002264.

PR 09-APR-1998; 98AU-00002911.
 PR 23-APR-1998; 98AU-00003128.
 PR 05-MAY-1998; 98AU-00003338.
 PR 22-MAY-1998; 98AU-00003654.
 PR 29-JUL-1998; 98AU-00004917.
 PR 30-JUL-1998; 98AU-00004963.
 PR 04-AUG-1998; 98AU-00005028.
 XX
 XX (CSLC-) CSL LTD.
 XX
 XX Ross BC, Barr IG, Patterson MA, Agius CT, Rothel LJ, Margetts MB;
 PI Hocking DM, Webb EA;
 XX
 DR WPI; 1999-385613/32.
 DR N-PSDB; AAX91618.
 XX
 PT Antigenic Porphyromonas gingivalis peptides for preventing gingivitis.
 XX
 XX Claim 1; Page 371; 588pp; English.
 XX
 CC AAX91536 to AAX91801 encode two hundred and sixty six antigenic
 CC Porphyromonas gingivalis (PG) polypeptide sequences given in AAY34318 to
 CC AAY34583. AAX91802 to AAX91989 represent PCR primers used in the
 CC isolation of the PG polypeptides. The PG polypeptides have antibacterial
 CC activity with a vaccine mechanism of action. The PG polypeptides can be
 CC used as vaccines especially against Porphyromonas gingivalis. Probes can
 CC be used to detect Porphyromonas gingivalis in standard hybridisation
 CC assays. Porphyromonas gingivalis is involved in periodontal disease
 CC especially gingivitis. (Updated on 20-MAR-2003 to correct PR field.)
 CC (Updated on 27-AUG-2003 to correct OS field.)
 XX
 SQ Sequence 221 AA;
 Query Match 81.8%; Score 18; DB 2; Length 221;
 Best Local Similarity 57.1%; Pred. NO. 2.9e+03;
 Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 VXXXXAEF 7
 Db 73 VGNAAEF 79
 RESULT 59
 AAY85820
 ID AAY85820 standard; protein; 225 AA.
 XX
 AC AAY85820;
 XX
 DT 10-APR-2000 (first entry)
 XX
 DE S. pneumoniae derived protein #29.
 XX
 KW Treatment; prevention; disease; diagnosis; gene therapy; screening;
 KW bacterial; antimicrobial; antibiotic; pathogenesis; infection.
 XX
 OS Streptococcus pneumoniae.
 XX
 PN WO9806734-A1.
 XX
 PD 19-FEB-1998.
 XX
 PF 15-AUG-1997; 97WO-US014436.
 XX
 PR 16-AUG-1996; 96US-0024022P.
 XX
 XX (SMIK) SMITHKLINE BEECHAM CORP.
 XX
 XX Black MT, Hodgson JE, Knowles DJC, Lonetto MA, Nicholas RO;
 PI Stodola RK;
 XX
 DR WPI; 1998-159452/14.
 DR N-PSDB; AAZ96199.
 XX

PT Streptococcus pneumoniae proteins and related DNA - useful for screening
PT compounds for antibacterial activity.
XX
PS Claim 5; Page 338-339; 640pp; English.
XX
CC This invention describes novel isolated Streptococcus pneumoniae
CC polynucleotides (see AAZ96173-286494) and their encoded proteins (see
CC AAY85792-Y86182). The DNA, vectors and host cells described in the method
CC of the invention are useful for the recombinant expression of the
CC polypeptides. The polypeptides are useful for treatment or prevention of
CC disease, or diagnosis of disease related to expression or activity of
CC such a polypeptide. They can also be used to screen for compounds which
CC interact with and inhibit or activate such a polypeptide. The
CC polypeptides (or DNA encoding them, via gene therapy) are also useful for
CC inducing an immunological response in a mammal. The antagonists are
CC useful to inhibit such bacterial polypeptides. The polypeptides are
CC particularly useful to identify antimicrobial compounds and antibiotics.
CC They are also useful to determine their role in pathogenesis of
CC infection, dysfunction and disease
XX
SQ Sequence 225 AA;

Query Match 81.8%; Score 18; DB 2; Length 225;
Best Local Similarity 57.1%; Pred. No. 3e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
| | | |
Db 115 VMSAAEF 121

RESULT 60
ABB65926
ID ABB65926 standard; protein; 225 AA.
XX
AC ABB65926;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 24570.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX
OS Drosophila melanogaster.
XX
XX WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US009231.
XX
PR 23-MAR-2000; 2000US-0191637P.
PR
PR 11-JUL-2000; 2000US-00614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li FWD, Myers EW;
XX
XX WPI; 2001-656860/75.
DR
DR N-PSDB; ABL10029.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.
PT
XX Disclosure; SEQ ID NO 24570; 21pp + Sequence Listing; English.
PS
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention

CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pt_sequences
XX
SQ Sequence 225 AA;

Query Match 81.8%; Score 18; DB 4; Length 225;
Best Local Similarity 57.1%; Pred. No. 3e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
| | | |
Db 7 VITAAEF 13

RESULT 61
ADJ49593
ID ADJ49593 standard; protein; 228 AA.
XX
AC ADJ49593;
XX
DT 06-MAY-2004 (first entry)
XX
DE Oil-associated gene related protein #1093.
XX
KW oil-associated gene; transgenic; enhanced seed oil; vegetable oil.
XX
OS Unidentified.
XX
PN US2004025202-A1.
XX
PD 05-FEB-2004.
XX
PF 14-MAR-2003; 2003US-00389566.
XX
PR 15-MAR-2002; 2002US-0365301P.
PR
PR 26-JUN-2002; 2002US-0391786P.
PR
PR 26-JUN-2002; 2002US-0392018P.
XX
PA (LAUR/) LAURIE C C.
PA (RAVA/) RAVANELLO M.
PA (SAVA/) SAVAGE T.
PA (LEDE/) LEDEAUX J R.
PA (ROGE/) ROGERS J A.
XX
PI Laurie CC, Ravanello M, Savage T, Ledeaux JR, Rogers JA;
XX
XX WPI; 2004-142683/14.
DR
XX Novel recombinant DNA construct comprising a promoter functional in
PT plants operably linked to an oil-associated gene for producing transgenic
PT plant seed.
XX
XX Example 3; SEQ ID NO 1597; 22pp; English.
PS
XX The invention relates to a recombinant DNA construct comprising a
CC promoter functional in plants operably linked to an oil-associated gene.
CC The construct is useful for transgenic plant seed which has in its genome
CC the construct, that is functional in the plant to transcribe the oil-
CC associated gene. The transgenic plant seed grows into a plant having
CC enhanced seed oil as compared to wild type. The construct is useful for
CC producing hybrid maize seed. The transgenic plant seed is useful for
CC producing vegetable oil. The present sequence represents the amino acid
CC sequence of an oil-associated gene related protein.
XX
XX Sequence 228 AA;

Query Match 81.8%; Score 18; DB 8; Length 228;
Best Local Similarity 57.1%; Pred. No. 3e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

QY 1 VXXAAEF 7
Db 41 VEAFAEF 47

RESULT 62
ADJ49577
ID ADJ49577 standard; protein; 228 AA.
XX
AC ADJ49577;
XX
DT 06-MAY-2004 (first entry)
XX
DE Oil-associated gene related protein #1077.
XX
DE oil-associated gene; transgenic; enhanced seed oil; vegetable oil.
XX
OS Unidentified.
XX
PN US2004025202-A1.
XX
PD 05-FEB-2004.
XX
PF 14-MAR-2003; 2003US-00389566.
XX
PR 15-MAR-2002; 2002US-0365301P.
XX
PR 26-JUN-2002; 2002US-0391786P.
XX
PR 26-JUN-2002; 2002US-0392018P.
XX
PA (LAUR/) LAURIE C C.
PA (RAVA/) RAVANELLO M.
PA (SAVA/) SAVAGE T.
PA (LEDE/) LEDEAUX J R.
PA (ROGE/) ROGERS J A.
XX
PI Laurie CC, Ravanello M, Savage T, Ledeaux JR, Rogers JA;
XX
PI WPI; 2004-142683/14.
XX
DR
XX
PT Novel recombinant DNA construct comprising a promoter functional in
PT plants operably linked to an oil-associated gene for producing transgenic
PT plant seed.
XX
PS Example 3; SEQ ID NO 1581; 22pp; English.
XX
CC The invention relates to a recombinant DNA construct comprising a
CC promoter functional in plants operably linked to an oil-associated gene.
CC The construct is useful for transgenic plant seed which has in its genome
CC the construct, that is functional in the plant to transcribe the oil-
CC associated gene. The transgenic plant seed grows into a plant having
CC enhanced seed oil as compared to wild type. The construct is useful for
CC producing hybrid maize seed. The transgenic plant seed is useful for
CC producing vegetable oil. The present sequence represents the amino acid
CC sequence of an oil-associated gene related protein.
XX
SQ Sequence 228 AA;

Query Match 81.8%; Score 18; DB 8; Length 228;
Best Local Similarity 57.1%; Pred. No. 3e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
Db 41 VEAFAEF 47

RESULT 63
AAW60988
ID AAW60988 standard; protein; 231 AA.
XX
AC AAW60988;
XX
DT 13-OCT-1998 (first entry)
XX

```

```

XX Streptococcus pneumoniae encoded polypeptide.
DE coding region; ORF; open reading frame; antibacterial; infection;
XX prevention; meningitis.
KW Streptococcus pneumoniae.
XX
OS WO9819689-A1.
XX
PN 14-MAY-1998.
XX
PD 27-OCT-1997; 97WO-US019226.
XX
PF 01-NOV-1996; 96US-0029930P.
XX
PR (SMIK ) SMITHKLINE BEECHAM CORP.
XX (SMIK ) SMITHKLINE BEECHAM PLC.
PA
XX
PI Black MT, Hodgson JE, Knowles DJC, Lonetto MA, Nicholas RO;
PI Reid RH, Zarfos PN;
XX
XX WPI; 1998-286586/25.
XX N-PSDB; AAV37380.
XX
PT New isolated nucleic acids from Streptococcus pneumoniae - useful, e.g.
PT for identifying anti-bacterial(s) for treatment and prevention of
PT meningitis.
XX
PS Claim 11; Page 85; 130pp; English.
XX
CC The sequence is that of the polypeptide encoded by a region isolated from
CC S. pneumoniae which shows homology to serine protease HtrA. The protein,
CC or agonists of it, may be useful as an antibacterial for treatment or
CC prevention of infection, specifically caused by S. pneumoniae
CC (particularly meningitis) but possibly also Helicobacter pylori (ulcers
CC and gastric cancer). It may be of particular use before insertion of an
CC in-dwelling device or any other invasive procedure. The protein, or
CC nucleic acid encoding it, can also be used in vaccines to induce a
CC cellular and/or humoral immune response, or to screen for other
CC antibacterials. The DNA may also contain flanking sequences that are
CC potential sources of control elements for bacterial gene expression.
CC Detecting a sequence encoding the protein can be used diagnostically,
CC e.g. to detect a mutation for serotyping or classifying infectious agents
XX
SQ Sequence 231 AA;

Query Match 81.8%; Score 18; DB 2; Length 231;
Best Local Similarity 57.1%; Pred. No. 3.1e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
Db 1 VTTVAEF 7

RESULT 64
AA99822
ID AA99822 standard; protein; 236 AA.
XX
AC AA99822;
XX
DT 19-SEP-2000 (first entry)
XX
DE Soybean partial protein phosphatase 2A regulatory subunit B.
XX
KW Soybean; protein phosphatase 2A; protein phosphorylation modulation;
XX transgenic plant; gene therapy.
XX
OS Glycine max.
XX
PN WO200036121-A2.
XX

```


PR 25-OCT-1999; 99US-0161404P.
 PR 25-OCT-1999; 99US-0161405P.
 PR 25-OCT-1999; 99US-0161406P.
 PR 26-OCT-1999; 99US-0161359P.
 PR 26-OCT-1999; 99US-0161360P.
 PR 26-OCT-1999; 99US-0161361P.
 PR 28-OCT-1999; 99US-0161920P.
 PR 28-OCT-1999; 99US-0161922P.
 PR 28-OCT-1999; 99US-0161933P.
 PR 29-OCT-1999; 99US-0162142P.

Query Match 81.8%; Score 18; DB 3; Length 239;
 Best Local Similarity 57.1%; Pred. No. 3.2e+03;
 Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
 | |||
 Db 191 VTTRAEF 197

RESULT 67
 AAR97195
 ID AAR97195 standard; protein; 240 AA.

AC AAR97195;
 XX
 DT 17-OCT-1996 (first entry)
 XX
 DE E.coli nitroreductase protein.
 XX
 KW E.coli; nfsA; nitroreductase; mutagen; carcinogen.
 XX
 OS Escherichia coli.

XX
 PN JP08131176-A.
 XX
 PD 28-MAY-1996.

XX
 PF 07-NOV-1994; 94JP-002989336.
 XX
 PR 07-NOV-1994; 94JP-002989336.

XX (CHCC) CHISSO CORP.
 XX
 DR WPI; 1996-303853/31.
 DR N-PSDB; AAT32986.

XX E. coli nitroreductase gene - useful to produce mutagen or carcinogen
 PT sensitive strain, using E. coli host.
 XX
 PS Claim 2; Page 3-4; 6pp; Japanese.

XX This is the amino acid sequence of the nitroreductase enzyme ncoded by
 CC the E.coli nfsA gene. The gene was isolated as given in Kumar and
 CC Jayaraman, (1991) J.Bioscience, 15, 145-159, and was subcloned to produce
 CC the plasmid pAJ102. The gene can be used to generate strains of E.coli
 CC which are sensitive to mutagenic and carcinogenic substances
 XX
 SQ Sequence 240 AA;

Query Match 81.8%; Score 18; DB 2; Length 240;
 Best Local Similarity 57.1%; Pred. No. 3.2e+03;
 Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
 | |||
 Db 70 VAQADEF 76

RESULT 68
 AAU29379
 ID AAU29379 standard; protein; 240 AA.

AC AAU29379;
 XX
 DT 18-DEC-2001 (first entry)
 XX
 DE Novel mar regulated protein (NIMR) #51.
 XX
 KW mar regulated polypeptide; NIMR; microbial infection; antibacterial.
 XX
 OS Escherichia coli.
 XX
 PN WO200170776-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 08-MAR-2001; 2001WO-US007478.
 PR 10-MAR-2000; 2000US-0188362P.
 XX
 PA (TUFT) TUFTS COLLEGE.

XX Levy SB, Barbosa TM, Alekshun MN;
 PI
 DR WPI; 2001-602769/68.
 DR N-PSDB; AAS46278.

XX Identifying compounds that modulate a newly identified mar regulated
 PT polypeptide activity, useful as antimicrobial compounds, involves
 PT contacting the polypeptide with a test compound.

XX Disclosure; Page 526; 526pp; English.

XX The invention relates to a method of identifying compounds that modulate
 CC a newly identified mar regulated (NIMR) polypeptide activity. The method
 CC comprises contacting an NIMR polypeptide with a test compound under
 CC interaction conditions, determining the ability of the compound to
 CC modulate the activity or expression of the polypeptide, and selecting the
 CC modulators. NIMR nucleic acids and polypeptides are used in the treatment
 CC of microbial infections, and in screening for modulators of NIMR
 CC expression and activity. These modulators can be used to reduce the
 CC infectivity of a microbe on a surface, and the virulence of a microbe in
 CC a subject suffering from an infection. AAU29329-AAU29379 represent
 CC Escherichia coli NIMR amino acid sequences of the invention

XX Sequence 240 AA;

Query Match 81.8%; Score 18; DB 4; Length 240;
 Best Local Similarity 57.1%; Pred. No. 3.2e+03;
 Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
 | |||
 Db 70 VAQADEF 76

RESULT 69
 ABP57429
 ID ABP57429 standard; protein; 240 AA.

XX
 AC ABP57429;

XX
 DT 24-APR-2003 (first entry)

XX E. coli major nitroreductase protein NfsA amino acid sequence.

XX Escherichia coli; major nitroreductase protein; NfsA; Vibrio harveyi;
 KW NADH oxidoreductase Frp; flavin reductase P; enzyme; antibacterial;
 KW antituberculostatic; vulnery; antiinflammatory; virucide; auditory;
 KW antitussive; litholytic; nephrotropic; gastrointestinal; antitumor;
 KW bacterial growth inhibitor; bacterial infection; microbial infection;
 KW tuberculosis; respiratory infection; bronchitis; sinusitis; pneumonia;
 KW sore throat; diphtheria; acute epiglottitis; influenza; otitis media;
 KW bronchopneumonia; Legionnaire's disease; whooping cough; urethritis;
 KW cystitis; pyelonephritis; colitis; enteritis; gastric ulcer; cholera;

KW duodenal ulcer; pancreatitis; typhus.
 XX Escherichia coli.
 OS WO2003000255-A1.
 XX 03-JAN-2003.
 PD 25-JUN-2002; 2002WO-US020387.
 PF 25-JUN-2001; 2001US-0300636P.
 XX (ESSE-) ESSENTIAL THERAPEUTICS INC.
 XX (PLIV) PLIVA DD ZAGREB.
 XX Magee AS, Roy A, Moe ST, Griffith JP, Ala PJ, Ali J, Clement JJ;
 XX Navia M;
 XX WPI; 2003-201374/19.
 DR New nitrofurantoin containing heterocyclic compounds useful in treating e.g.
 PT tuberculosis.
 XX Disclosure; Page 23; 50pp; English.
 XX The present invention describes nitrofurantoin compounds substituted with
 CC heterocyclic compounds (I) or their salts. (I) has antibacterial,
 CC antituberculostatic, vulnerary, antiinflammatory, virucide, auditory,
 CC antitussive, litholytic, nephrotropic, gastrointestinal and antiulcer
 CC activities, and can be used as a bacterial growth inhibitor. (I) can be
 CC used for treating bacterial or microbial infection and tuberculosis in
 CC humans as well as animals. (I) can also be used as sterilizants,
 CC antiseptics, adjuvants in wound dressings (e.g. bandages) and adjuvants
 CC in wound cleaning methods (e.g. swipes or gavage); for treating upper and
 CC lower respiratory tract infections (e.g. bronchitis, sinusitis,
 CC pneumonia, sore throat, chronic streptococcal infections, diphtheria,
 CC acute epiglottitis, influenza, chronic bronchitis, middle ear infections
 CC (otitis media), bronchopneumonia, Legionnaire's disease, atypical
 CC pneumonia or whooping cough), urinary tract infections (e.g. urethritis,
 CC cystitis, pyelonephritis (kidney infection), asymptomatic bacteriuria,
 CC interstitial cystitis, acute urethral syndrome or recurrent urinary tract
 CC infections), gastrointestinal infections of bacterial origin (e.g. food
 CC borne infections, colitis, enteritis, gastric ulcers, duodenal ulcers,
 CC pancreatitis, gall bladder infections, cholera or typhus). (I) can be
 CC used for inhibiting bacterial growth in a non-living system. The present
 CC sequence represents an E. coli major nitroreductase protein NfsA which is
 CC given in comparison with a NADH oxidoreductase Fp (flavin reductase P)
 CC amino acid sequence from the luminescent bacterium Vibrio harveyi in the
 CC exemplification of the present invention
 XX Sequence 240 AA;
 SQ
 Query Match 81.8%; Score 18; DB 6; Length 240;
 Best Local Similarity 57.1%; Pred. No. 3.2e+03;
 Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 VXXAAEF 7
 Db 70 VESAAEF 76
 RESULT 70
 ABP57430
 ID ABP57430 standard; protein; 240 AA.
 XX
 AC ABP57430;
 XX
 DT 24-APR-2003 (first entry)
 XX
 DE NADH oxidoreductase Fp (flavin reductase P) amino acid sequence.
 XX
 KW Escherichia coli; major nitroreductase protein; NfsA; Vibrio harveyi;
 KW NADH oxidoreductase Fp; flavin reductase P; enzyme; antibacterial;

KW antituberculostatic; vulnerary; antiinflammatory; virucide; auditory;
 KW antitussive; litholytic; nephrotropic; gastrointestinal; antiulcer;
 KW bacterial growth inhibitor; bacterial infection; microbial infection;
 KW tuberculosis; respiratory infection; bronchitis; sinusitis; pneumonia;
 KW sore throat; diphtheria; acute epiglottitis; influenza; otitis media;
 KW bronchopneumonia; Legionnaire's disease; whooping cough; urethritis;
 KW cystitis; pyelonephritis; colitis; enteritis; gastric ulcer; cholera;
 KW duodenal ulcer; pancreatitis; typhus.
 OS Vibrio harveyi.
 XX WO2003000255-A1.
 XX 03-JAN-2003.
 XX 25-JUN-2002; 2002WO-US020387.
 XX 25-JUN-2001; 2001US-0300636P.
 XX (ESSE-) ESSENTIAL THERAPEUTICS INC.
 XX (PLIV) PLIVA DD ZAGREB.
 XX Magee AS, Roy A, Moe ST, Griffith JP, Ala PJ, Ali J, Clement JJ;
 XX Navia M;
 XX WPI; 2003-201374/19.
 DR New nitrofurantoin containing heterocyclic compounds useful in treating e.g.
 PT tuberculosis.
 XX Disclosure; Page 23; 50pp; English.
 XX The present invention describes nitrofurantoin compounds substituted with
 CC heterocyclic compounds (I) or their salts. (I) has antibacterial,
 CC antituberculostatic, vulnerary, antiinflammatory, virucide, auditory,
 CC antitussive, litholytic, nephrotropic, gastrointestinal and antiulcer
 CC activities, and can be used as a bacterial growth inhibitor. (I) can be
 CC used for treating bacterial or microbial infection and tuberculosis in
 CC humans as well as animals. (I) can also be used as sterilizants,
 CC antiseptics, adjuvants in wound dressings (e.g. bandages) and adjuvants
 CC in wound cleaning methods (e.g. swipes or gavage); for treating upper and
 CC lower respiratory tract infections (e.g. bronchitis, sinusitis,
 CC pneumonia, sore throat, chronic streptococcal infections, diphtheria,
 CC acute epiglottitis, influenza, chronic bronchitis, middle ear infections
 CC (otitis media), bronchopneumonia, Legionnaire's disease, atypical
 CC pneumonia or whooping cough), urinary tract infections (e.g. urethritis,
 CC cystitis, pyelonephritis (kidney infection), asymptomatic bacteriuria,
 CC interstitial cystitis, acute urethral syndrome or recurrent urinary tract
 CC infections), gastrointestinal infections of bacterial origin (e.g. food
 CC borne infections, colitis, enteritis, gastric ulcers, duodenal ulcers,
 CC pancreatitis, gall bladder infections, cholera or typhus). (I) can be
 CC used for inhibiting bacterial growth in a non-living system. The present
 CC sequence represents an E. coli major nitroreductase protein NfsA which is
 CC given in comparison with a NADH oxidoreductase Fp (flavin reductase P)
 CC amino acid sequence from the luminescent bacterium Vibrio harveyi in the
 CC exemplification of the present invention
 XX Sequence 240 AA;
 SQ
 Query Match 81.8%; Score 18; DB 6; Length 240;
 Best Local Similarity 57.1%; Pred. No. 3.2e+03;
 Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 VXXAAEF 7
 Db 70 VESAAEF 76
 RESULT 71
 AAY34534
 ID AAY34534 standard; protein; 243 AA.
 XX
 AC AAY34534;

XX DT 27-AUG-2003 (revised)
 DT 20-MAR-2003 (revised)
 DT 25-AUG-1999 (first entry)
 XX DE Porphorymonas gingivalis protein PG65.
 XX KW Porphorymonas gingivalis; PG; periodontal disease; gingivitis; vaccine;
 KW antigenic.
 XX OS Porphorymonas gingivalis.
 XX PN WO9929870-A1.
 XX PD 17-JUN-1999.
 XX PF 10-DEC-1998; 98WO-AU001023.
 XX PR 10-DEC-1997; 97AU-00000839.
 XX PR 31-DEC-1997; 97AU-00001182.
 XX PR 30-JAN-1998; 98AU-00001546.
 XX PR 10-MAR-1998; 98AU-00002264.
 XX PR 09-APR-1998; 98AU-00002911.
 XX PR 23-APR-1998; 98AU-00003128.
 XX PR 05-MAY-1998; 98AU-00003338.
 XX PR 22-MAY-1998; 98AU-00003654.
 XX PR 29-JUN-1998; 98AU-00004917.
 XX PR 30-JUL-1998; 98AU-00004963.
 XX PR 04-AUG-1998; 98AU-00005028.
 XX PA (CSLC-) CSL LTD.
 XX PI Ross BC, Barr IG, Patterson MA, Agius CT, Rothel LJ, Margette MB;
 PI Hocking DM, Webb EA;
 XX WPI; 1999-385613/32.
 DR N-PSDB; AAX91752.
 XX NT Antigenic Porphorymonas gingivalis peptides for preventing gingivitis.
 XX PS Claim 1; Page 524-525; 588pp; English.
 XX CC AAX91536 to AAX91801 encode two hundred and sixty six antigenic
 CC Porphorymonas gingivalis (PG) polypeptide sequences given in AAX94318 to
 CC AAX94383. AAX91802 to AAX91989 represent PCR primers used in the
 CC isolation of the PG polypeptides. The PG polypeptides have antibacterial
 CC activity with a vaccine mechanism of action. The PG polypeptides can be
 CC used as vaccines especially against Porphorymonas gingivalis. Probes can
 CC be used to detect Porphorymonas gingivalis in standard hybridisation
 CC assays. Porphorymonas gingivalis is involved in periodontal disease
 CC especially gingivitis. (Updated on 20-MAR-2003 to correct PR field.)
 CC (Updated on 27-AUG-2003 to correct OS field.)
 XX SQ Sequence 243 AA;
 Query Match 81.8%; Score 18; DB 2; Length 243;
 Best Local Similarity 57.1%; Pred. No. 3.2e+03;
 Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 VXXAAEF 7
 Db 56 VGAAAEF 62
 RESULT 72
 AAU67035
 ID AAU67035 standard; protein; 247 AA.
 XX AC AAU67035;
 XX DT 27-FEB-2002 (first entry)
 XX DE Porphionibacterium acnes immunogenic protein #27931.
 XX KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
 KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 KW dermatological; osteopathic; neuroprotectant.
 XX OS Propionibacterium acnes.
 XX PN WO200181581-A2.
 XX PD 01-NOV-2001.
 XX PF 20-APR-2001; 2001WO-US012865.
 XX PR 21-APR-2000; 2000US-0199047P.
 XX PR 02-JUN-2000; 2000US-0208841P.
 XX PR 07-JUL-2000; 2000US-0216747P.
 XX PA (CORI-) CORIXA CORP.
 XX PI Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
 PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
 XX WPI; 2001-616774/71.
 DR N-PSDB; AAS9774.
 XX PT Propionibacterium acnes polypeptides and nucleic acids useful for
 PT vaccinating against and diagnosing infections, especially useful for
 PT treating acne vulgaris.
 XX PS Example 1; SEQ ID NO 28230; 1069pp; English.
 XX CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
 CC polypeptides. The proteins and their associated DNA sequences are used in
 CC the treatment, prevention and diagnosis of medical conditions caused by
 CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
 CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
 CC P. acnes is also involved in infections of bone, joints and the central
 CC nervous system, however it is particularly involved in the inflammatory
 CC lesions associated with acne vulgaris. A method for detecting the
 CC presence or absence of P. acnes in a patient comprises contacting a
 CC sample with a binding agent that binds to the proteins of the invention
 CC and determining the amount of bound protein in the sample. The
 CC polypeptides may be used as antigens in the production of antibodies
 CC specific for P. acnes proteins. These antibodies can be used to
 CC downregulate expression and activity of P. acnes polypeptides and
 CC therefore treat P. acnes infections. The antibodies may also be used as
 CC diagnostic agents for determining P. acnes presence, for example, by
 CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
 CC this patent did not form part of the printed specification, but was
 CC obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX SQ Sequence 247 AA;
 Query Match 81.8%; Score 18; DB 4; Length 247;
 Best Local Similarity 57.1%; Pred. No. 3.3e+03;
 Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 VXXAAEF 7
 Db 57 VAVTAEF 63
 RESULT 73
 ABM63554
 ID ABM63554 standard; protein; 247 AA.
 XX AC ABM63554;
 XX DT 20-OCT-2003 (first entry)
 XX DE Propionibacterium acnes predicted ORF-encoded polypeptide #28230.

XX Acne vulgaris; antiseborrheic; dermatological; antibacterial;
 KW immunostimulant; immune response; vaccine.
 XX Propionibacterium acnes.
 XX WO2003033515-A1.
 XX 24-APR-2003.
 XX 11-OCT-2002; 2002WO-US032727.
 XX 15-OCT-2001; 2001US-00978825.
 XX (CORI-) CORIXA CORP.
 XX Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;
 PI Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;
 PI Barth B, Vallie-Douglass J;
 XX WPI: 2003-381789/36.
 XX N-PSDB; ACF64703.
 XX New Propionibacterium acnes polypeptides and polynucleotides encoding the
 PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
 PT or for stimulating an immune response specific for a P. acnes protein.
 XX Example 1; SEQ ID NO 28230; 1481pp; English.
 XX The invention relates to an isolated polynucleotide (ACF64435-ACF64733)
 CC encoding a Propionibacterium acnes protein. The invention also relates to
 CC polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to
 CC immunogenic fragments of P. acnes polypeptides. The invention
 CC additionally encompasses expression vectors and host cells comprising a
 CC polynucleotide of the invention; antibodies against polypeptides of the
 CC invention; fusion proteins comprising a polypeptide of the invention; a
 CC method for stimulating an immune response specific for a P. acnes
 CC polypeptide and an isolated T cell population comprising T cells prepared
 CC via this method; a vaccine composition (comprising P. acnes polypeptides,
 CC polynucleotides, antibodies, fusion proteins, T cell populations, or
 CC antigen-presenting cells that express the polypeptide); a method and kit
 CC for detecting or determining the presence or absence of P. acnes in a
 CC patient; and a method for inhibiting the development of P. acnes in a
 CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion
 CC proteins, T cell populations or antigen-presenting cells that express the
 CC polypeptides are useful for diagnosing, preventing or treating acne
 CC vulgaris, or for stimulating an immune response specific for a P. acnes
 CC protein. The polynucleotides can also be used as probes or primers for
 CC nucleic acid hybridisation. The vaccine composition is useful for the
 CC stimulation of an immune response against P. acnes, or for treating acne,
 CC and the kit is useful for performing a diagnostic assay. The present
 CC sequence represents a polypeptide predicted to be encoded by an ORF (open
 CC reading frame) contained within the P. acnes polynucleotides of the
 CC invention. Note: The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 247 AA;
 Query Match 81.8%; Score 18; DB 6; Length 247;
 Best Local Similarity 57.1%; Pred. No. 3.3e+03;
 Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 VXXAAEF 7
 DB 57 VAVTAEF 63
 RESULT 74
 AAY34401
 ID AAY34401 standard; protein; 248 AA.
 XX
 AC AAY34401;

XX 27-AUG-2003 (revised)
 DT 20-MAR-2003 (revised)
 DT 25-AUG-1999 (first entry)
 XX
 DE Porphyromonas gingivalis protein PG65.
 XX
 KW Porphyromonas gingivalis; PG; periodontal disease; gingivitis; vaccine;
 KW antigenic.
 XX
 OS Porphyromonas gingivalis.
 XX WO9929870-A1.
 PN 17-JUN-1999.
 PD
 XX 10-DEC-1998; 98WO-AU001023.
 PF 10-DEC-1997; 97AU-00000839.
 XX 31-DEC-1997; 97AU-00001182.
 PR 30-JAN-1998; 98AU-00001546.
 PR 10-MAR-1998; 98AU-00002264.
 PR 09-APR-1998; 98AU-00002911.
 PR 23-APR-1998; 98AU-00003128.
 PR 05-MAY-1998; 98AU-00003338.
 PR 23-MAY-1998; 98AU-00003654.
 PR 29-JUL-1998; 98AU-00004917.
 PR 30-JUL-1998; 98AU-00004963.
 PR 04-AUG-1998; 98AU-00005028.
 XX (CSLC-) CSL LTD.
 XX Ross BC, Barr IG, Patterson MA, Agius CT, Rothel LJ, Margetts MB;
 PI Hocking DM, Webb EA;
 XX WPI: 1999-385613/32.
 DR N-PSDB; AAX91619.
 XX Antigenic Porphyromonas gingivalis peptides for preventing gingivitis.
 PT
 XX Claim 1; Page 372; 589pp; English.
 XX AAX91536 to AAX91801 encode two hundred and sixty six antigenic
 CC Porphyromonas gingivalis (PG) polypeptide sequences given in AAY34318 to
 CC AAY34583. AAX91802 to AAX91989 represent PCR primers used in the
 CC isolation of the PG polypeptides. The PG polypeptides have antibacterial
 CC activity with a vaccine mechanism of action. The PG polypeptides can be
 CC used as vaccines especially against Porphyromonas gingivalis. Probes can
 CC be used to detect Porphyromonas gingivalis in standard hybridisation
 CC assays. Porphyromonas gingivalis is involved in periodontal disease
 CC especially gingivitis. (Updated on 20-MAR-2003 to correct PR field.)
 CC (Updated on 27-AUG-2003 to correct OS field.)
 XX
 SQ Sequence 248 AA;
 Query Match 81.8%; Score 18; DB 2; Length 248;
 Best Local Similarity 57.1%; Pred. No. 3.3e+03;
 Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 VXXAAEF 7
 DB 61 VGAAAEF 67
 RESULT 75
 ABB69168
 ID ABB69168 standard; protein; 249 AA.
 XX
 AC ABB69168;
 XX 26-MAR-2002 (first entry)
 DT
 XX Drosophila melanogaster polypeptide SEQ ID NO 34296.
 DE

```
XX Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
KW Drosophila melanogaster.
OS
OS
PN WO200171042-A2.
XX
XX 27-SEP-2001.
XX
XX 23-MAR-2001; 2001WO-US009231.
XX
XX 23-MAR-2000; 2000US-0191637P.
PR
PR 11-JUL-2000; 2000US-00614150.
XX
XX (PEKE ) PE CORP NY.
PA
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
PI
XX WPI; 2001-656860/75.
DR N-PSDB; ABL13271.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.
XX
XX Disclosure; SEQ ID NO 34296; 21pp + Sequence Listing; English.
PS
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 249 AA;
SQ
Query Match 81.8%; Score 18; DB 4; Length 249;
Best Local Similarity 42.9%; Pred. No. 3.3e+03;
Matches 3; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 1 VXXAAEF 7
Db : |||
47 ISAAEF 53
RESULT 76
ADRO1255
ID ADR01255 standard; protein; 253 AA.
XX
XX ADR01255;
AC
XX
XX 21-OCT-2004 (first entry)
DT
DE Farnesyl dibenzodiazepinone biosynthetic ORF23 protein HYDK, SEQ ID 46.
XX
XX Cytostatic; Antibacterial; Antiinflammatory;
KW Cancer cell growth inhibitor; Lipooxygenase inhibitor;
KW farnesyl dibenzodiazepinone; cancer; bacterial infections; inflammation.
XX
XX Micromonospora sp.; strain 046-EC011.
OS
XX
XX WO2004065591-A1.
PN
XX
XX 05-AUG-2004.
PD
XX
XX 21-JAN-2004; 2004WO-CA000069.
XX
XX 21-JAN-2003; 2003US-0441126P.
PR
07-AUG-2003; 2003US-0492997P.
PR 10-NOV-2003; 2003US-0518286P.
XX
XX (ECOP-) ECOPIA BIOSCIENCES INC.
XX
XX Bachmann BO, Mcalpine JB, Zaropoulos E, Farnet CM, Piraeae M;
PI
XX WPI; 2004-593481/57.
DR N-PSDB; ADR01210, ADR01256.
XX
XX New farnesyl dibenzodiazepinone compounds useful for the treatment of
PT e.g. cancer, bacterial infections and inflammation.
XX
XX Claim 53; SEQ ID NO 46; 269pp; English.
XX
XX The present invention relates to novel farnesyl dibenzodiazepinone
CC compounds. The compounds can be obtained by cultivation of a novel strain
CC of Micromonospora sp., i.e. 046-EC011 or [S01]046. The compounds are
CC excellent tumour growth inhibitors as well mammalian lipooxygenase
CC inhibitors and so are useful for treating cancer, pre-cancerous
CC conditions, bacterial infections and inflammation in mammal. The present
CC sequence is a protein involved in the biosynthetic pathway for a farnesyl
CC dibenzodiazepinone.
XX
XX Sequence 253 AA;
SQ
Query Match 81.8%; Score 18; DB 8; Length 253;
Best Local Similarity 57.1%; Pred. No. 3.4e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 VXXAAEF 7
Db 123 VAVAAEF 129
RESULT 77
ABO63928
ID ABO63928 standard; protein; 255 AA.
XX
XX ABO63928;
AC
XX
XX 29-JUL-2004 (first entry)
DT
DE Klebsiella pneumoniae polypeptide seqid 10445.
XX
XX Recombinant expression vector; transcription regulatory element;
KW Klebsiella pneumoniae protein; antibacterial; vaccine.
XX
XX Klebsiella pneumoniae.
OS
XX
XX US6610836-B1.
PN
XX
XX 26-AUG-2003.
PD
XX
XX 27-JAN-2000; 2000US-00489039.
PF
XX
XX 29-JAN-1999; 99US-0117747P.
PR
XX
XX (GENO-) GENOME THERAPEUTICS CORP.
PA
XX
XX Breton GL, Osborne M;
PI
XX WPI; 2003-895346/82.
DR N-PSDB; ACH97479.
XX
XX New nucleic acid encoding a Klebsiella pneumoniae polypeptide, useful for
PT preparing a vaccine composition against Klebsiella pneumoniae.
XX
XX Disclosure; SEQ ID NO 10445; 932pp; English.
PS
XX
XX The invention describes a new isolated nucleic acid encoding a Klebsiella
CC pneumoniae polypeptide. Also described are: a recombinant expression
CC vector comprising the nucleic acid, operably linked to a transcription
```

CC regulatory element; and a cell comprising the recombinant expression
 CC vector. The nucleic acid is useful for preparing a vaccine composition
 CC against *Klebsiella pneumoniae*. This is the amino acid sequence of a
 CC *Klebsiella pneumoniae* polypeptide of the invention
 XX
 SQ Sequence 255 AA;

Query Match 81.8%; Score 18; DB 7; Length 255;
 Best Local Similarity 57.1%; Pred. No. 3.4e+03;
 Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VXXAAEF 7
 | |||
 Db 85 VAQAAEF 91

RESULT 78
 AAB79542
 ID AAB79542 standard; protein; 258 AA.

XX AAB79542;
 AC
 DT 30-APR-2001 (first entry)
 XX
 DE Corynebacterium glutamicum SMP protein sequence SEQ ID NO:600.
 XX
 KW Corynebacterium glutamicum; carbon metabolism and energy production;
 KW SMP protein; sugar metabolism and oxidative phosphorylation protein;
 KW fine chemical production; organic acid; proteinogenic amino acid;
 KW nonproteinogenic amino acid; purine base; pyrimidine base; nucleoside;
 KW nucleotide; lipid; saturated fatty acid; unsaturated fatty acid; diol;
 KW carbohydrate; aromatic compound; vitamin; cofactor; polyketide; enzyme;
 KW diagnosis; Corynebacterium diphtheriae; evolutionary study.
 XX

OS Corynebacterium glutamicum.

XX
 PN WO200100844-A2.

XX
 PD 04-JAN-2001.

XX
 PF 23-JUN-2000; 2000WO-IB000943.

XX
 PR 25-JUN-1999; 99US-0141031P.
 PR 08-JUL-1999; 99DE-01031412.
 PR 08-JUL-1999; 99DE-01031413.
 PR 08-JUL-1999; 99DE-01031419.
 PR 08-JUL-1999; 99DE-01031420.
 PR 08-JUL-1999; 99DE-01031424.
 PR 08-JUL-1999; 99DE-01031428.
 PR 08-JUL-1999; 99DE-01031431.
 PR 08-JUL-1999; 99DE-01031433.
 PR 08-JUL-1999; 99DE-01031434.
 PR 08-JUL-1999; 99DE-01031510.
 PR 08-JUL-1999; 99DE-01031562.
 PR 08-JUL-1999; 99DE-01031634.
 PR 09-JUL-1999; 99DE-01032180.
 PR 09-JUL-1999; 99DE-01032227.
 PR 09-JUL-1999; 99DE-01032230.
 PR 09-JUL-1999; 99US-0143208P.
 PR 14-JUL-1999; 99DE-01032924.
 PR 14-JUL-1999; 99DE-01032973.
 PR 14-JUL-1999; 99DE-01033005.
 PR 27-AUG-1999; 99DE-01040765.
 PR 31-AUG-1999; 99US-0151572P.
 PR 03-SEP-1999; 99DE-01042076.
 PR 03-SEP-1999; 99DE-01042079.
 PR 03-SEP-1999; 99DE-01042086.
 PR 03-SEP-1999; 99DE-01042087.
 PR 03-SEP-1999; 99DE-01042088.
 PR 03-SEP-1999; 99DE-01042095.
 PR 03-SEP-1999; 99DE-01042123.
 PR 03-SEP-1999; 99DE-01042125.

PA (BADI) BASF AG.

XX Pompejus M, Kroegeer B, Schroeder H, Zelder O, Haberhauer G;
 XX

DR WPI; 2001-061975/07.

DR N-PSDB; AAF71659.

XX
 PT New isolated Corynebacterium glutamicum nucleic acid encoding a sugar
 PT metabolism and oxidative phosphorylation protein for production or
 PT modulation of production of fine chemicals e.g. amino acids,
 PT carbohydrates or enzymes.

XX
 PS Claim 20; Page 1004-1005; 1246pp; English.

XX
 CC AAF71360 to AAF71750 encode the Corynebacterium glutamicum sugar
 CC metabolism and oxidative phosphorylation (SMP) proteins given in AAB79243
 CC to AAB 79633 which are involved in carbon metabolism and energy
 CC production. The C. glutamicum SMP gene can be used in vectors (II) for
 CC expression in host cells and production or modulation of production of
 CC fine chemicals, such as, an organic acid, a proteinogenic or
 CC nonproteinogenic amino acid (preferred), a purine or pyrimidine base, a
 CC nucleoside, a nucleotide, a lipid, a saturated or unsaturated fatty acid,
 CC a diol, a carbohydrate, an aromatic compound, a vitamin, a cofactor, a
 CC polyketide, or an enzyme. The presence of (I) or SMP proteins (III)
 CC encoded by them are used for diagnosing the presence or activity of
 CC Corynebacterium diphtheriae in a subject. (I), (II), (III) or host cells
 CC containing them are used to map genomes of organisms related to C.
 CC glutamicum, identify and localise C. glutamicum sequences of interest, in
 CC evolutionary studies, in determining SMP protein regions required for
 CC function, in modulating SMP protein activity, in modulating the
 CC metabolism of sugars, and in modulating high-energy molecule production
 CC in a cell (i.e. Atp, NADPH)

XX
 SQ Sequence 258 AA;

Query Match 81.8%; Score 18; DB 4; Length 258;
 Best Local Similarity 57.1%; Pred. No. 3.5e+03;
 Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VXXAAEF 7
 | |||
 Db 46 VATGAEF 52

RESULT 79

AAG92053

ID AAG92053 standard; protein; 258 AA.

XX
 AC AAG92053;

XX
 DT 26-SEP-2001 (first entry)

XX
 DE C glutamicum protein fragment SEQ ID NO: 5807.

XX
 KW Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
 KW organic acid synthesis.

XX
 OS Corynebacterium glutamicum.

XX
 PN EF1108790-A2.

XX
 PD 20-JUN-2001.

XX
 PF 18-DEC-2000; 2000EP-00127688.

XX
 PR 16-DEC-1999; 99JP-00377484.

XX
 PR 07-APR-2000; 2000JP-00159162.

XX
 PR 03-AUG-2000; 2000JP-00280988.

XX
 PA (KYOW) KYOWA HAKKO KOGYO KK.

XX
 PI Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
 PI Tateishi N, Senoh A, Ikeda M, Ozaki A;

XX WPI; 2001-376931/40.
DR N-PSDB; AAH67272.

XX Novel polynucleotides derived from Corynebacterium bacteria, for identifying
PT mutation point of a gene, measuring expression of a gene, analyzing
PT expression profile or pattern of a gene and identifying homologous gene.

XX Claim 17; SEQ ID NO 5807; 246pp + Sequence Listing; English.

XX The present invention provides a number of nucleotide and protein
CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These
CC are useful for identifying the mutation point of a gene derived from a
CC mutant of coryneform bacterium, measuring expression amount and analysing
CC the expression profile or expression pattern of a gene derived from
CC Coryneform bacterium, and identifying a homologue of a gene derived from
CC Coryneform bacterium. Coryneform bacteria are useful for producing amino
CC acids, nucleic acids, vitamins, saccharides and organic acids,
CC particularly L-lysine. The present sequence is a protein described in the
CC exemplification of the invention. Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from the European Patent Office

XX SQ Sequence 258 AA;

Query Match 81.8%; Score 18; DB 4; Length 258;
Best Local Similarity 57.1%; Pred. No. 3.5e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VXXXXAEF 7
| ||||
Db 46 VATGAEF 52

RESULT 80
AAV75744
ID AAV75744 standard; protein; 266 AA.

XX AC AAV75744;
XX 21-MAR-2000 (first entry)

XX DE Neisseria meningitidis ORF 981 protein sequence SEQ ID NO:2960.

XX KW Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
KW antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;
KW antibacterial; gene therapy.

XX OS Neisseria meningitidis.
XX WO9957280-A2.
XX PN 11-NOV-1999.
XX PD 30-APR-1999; 99WO-US009346.
XX PF 01-MAY-1998; 98US-0083758P.
XX PR 31-JUL-1998; 98US-0094869P.
XX PR 02-SEP-1998; 98US-0098994P.
XX PR 02-SEP-1998; 98US-0099062P.
XX PR 09-OCT-1998; 98US-0103749P.
XX PR 09-OCT-1998; 98US-0103794P.
XX PR 25-FEB-1999; 99US-0121528P.

XX PA (CHIR) CHIRON CORP.
XX (GENO-) INST GENOMIC RES.

XX PI Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;
PI Petersen J, Pizzo M, Rappuoli R, Ratti G, Scalato E, Scarselli M;
PI Tettelin H, Venter JC;
XX WPI; 2000-062150/05.

XX DR N-PSDB; AAH67272.

PT Novel Neisserial polypeptides predicted to be useful antigens for
PT vaccines and diagnostics.

XX
XX
PS Claim 2; Page 1383; 1453pp; English.

XX AAZ53015 to AAZ54536, AAZ54577 to AAZ54615, and AAZ74253 to AAZ75941
CC represent novel *Neisseria meningitidis* and *N. gonorrhoeae* polynucleotides
CC and polypeptides. AAZ54537 to AAZ54576 and AAZ54616 to AAZ55473 represent
CC PCR primers used in the exemplification of the present invention. The
CC polypeptides, the polynucleotides, antibodies and compositions of the
CC invention can be used as vaccines, as diagnostic reagents, and as
CC immunogenic compositions. The polypeptides can be used in the manufacture
CC of medicaments for treating or preventing infection due to *Neisseria*
CC bacteria (e.g. meningitis and septicaemia), to detect the presence of
CC *Neisseria* bacteria, or to raise antibodies. They may also be used to
CC screen for agonists or antagonists, which may themselves have use as
CC antibacterial agents. The polynucleotides of the invention may also be
CC used in gene therapy protocols
XX
SQ Sequence 266 AA;

Query Match 81.8%; Score 18; DB 3; Length 266;
Best Local Similarity 57.1%; Pred. No. 3.6e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
| | | |
Db 39 VASNAEF 45

RESULT 82

AAZ75742
ID AAZ75742 standard; protein; 266 AA.

XX
AC AAZ75742;

XX 12-SEP-2003 (revised)
DT 21-MAR-2000 (first entry)

XX *Neisseria gonorrhoeae* ORF 981 protein sequence SEQ ID NO:2956.

XX *Neisseria meningitidis*; *Neisseria gonorrhoeae*; antigen; vaccine;
KW antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;
KW antibacterial; gene therapy.

XX *Neisseria gonorrhoeae*.

XX WO9957280-A2.

XX 11-NOV-1999.

XX 30-APR-1999; 99WO-US0009346.

XX 01-MAY-1998; 98US-0083758P.

XX 31-JUL-1998; 98US-0094869P.

XX 02-SEP-1998; 98US-0098994P.

XX 02-SEP-1998; 98US-0099062P.

XX 03-OCT-1998; 98US-0103749P.

XX 09-OCT-1998; 98US-0103794P.

XX 09-OCT-1998; 98US-0103796P.

XX 25-FEB-1999; 99US-0121528P.

XX (CHIR) CHIRON CORP.

PA (GENO-) INST GENOMIC RES.

XX Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;
PI Petersen J, Pizsa M, Rappuoli R, Ratti G, Scalato E, Scarselli M;
PI Tettelin H, Venter JC;

XX WPI; 2000-062150/05.

DR N-PSDB; AAZ54504.

PT Novel *Neisseria* polypeptides predicted to be useful antigens for

PT vaccines and diagnostics.

XX
XX Claim 2; Page 1383; 1453pp; English.

XX AAZ53015 to AAZ54536, AAZ54577 to AAZ54615, and AAZ74253 to AAZ75941
CC represent novel *Neisseria meningitidis* and *N. gonorrhoeae* polynucleotides
CC and polypeptides. AAZ54537 to AAZ54576 and AAZ54616 to AAZ55473 represent
CC PCR primers used in the exemplification of the present invention. The
CC polypeptides, the polynucleotides, antibodies and compositions of the
CC invention can be used as vaccines, as diagnostic reagents, and as
CC immunogenic compositions. The polypeptides can be used in the manufacture
CC of medicaments for treating or preventing infection due to *Neisseria*
CC bacteria (e.g. meningitis and septicaemia), to detect the presence of
CC *Neisseria* bacteria, or to raise antibodies. They may also be used to
CC screen for agonists or antagonists, which may themselves have use as
CC antibacterial agents. The polynucleotides of the invention may also be
CC used in gene therapy protocols. (Updated on 12-SEP-2003 to standardise OS
CC field)
XX
SQ Sequence 266 AA;

Query Match 81.8%; Score 18; DB 3; Length 266;
Best Local Similarity 57.1%; Pred. No. 3.6e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
| | | |
Db 39 VASNAEF 45

RESULT 83

AAZ95653
ID AAZ95653 standard; protein; 268 AA.

XX
AC AAZ95653;

XX 25-OCT-2000 (first entry)

XX *Neisseria meningitidis* BASB071 gene-encoded protein.

DE *Neisseria meningitidis* BASB071 gene-encoded protein.

XX BASB071 gene; infection; diagnosis; therapy; vaccine.

XX *Neisseria meningitidis*.

XX WO200042191-A2.

XX 20-JUL-2000.

XX 10-JAN-2000; 2000WO-EP000135.

XX 15-JAN-1999; 99GB-00000838.

XX 15-JAN-1999; 99GB-00000952.

XX 28-JAN-1999; 99GB-00001945.

XX 28-JAN-1999; 99GB-00001948.

XX 29-JAN-1999; 99GB-00002074.

XX 29-JAN-1999; 99GB-00002078.

XX 29-JAN-1999; 99GB-00002088.

XX 09-FEB-1999; 99GB-00002879.

XX 10-FEB-1999; 99GB-00002936.

XX 20-FEB-1999; 99GB-00003978.

XX 23-FEB-1999; 99GB-00004133.

XX 25-FEB-1999; 99GB-00004404.

XX (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.

XX Ruelle J, Thonnard J;

XX WPI; 2000-476061/41.

XX N-PSDB; AA50077.

XX New *Neisseria meningitidis* BASB051, BASB057, BASB060, BASB061, BASB063,
PT BASB065, BASB066 and BASB071 polynucleotides and polypeptides used in the
PT treatment and diagnosis of *N. meningitidis* disease.

XX PS Claim 1; Page 110; 127pp; English.

XX CC The present sequence is that of the protein product of the BASB071 gene (see AAA50077) of *Neisseria meningitidis* ATCC 13090. The protein contains a leader sequence characteristic of a lipoprotein signal sequence, and shows significant similarity to the *Neisseria gonorrhoeae* HsdJ protein.

XX CC The invention provides novel *N. meningitidis* BASB051, BASB057, BASB060, BASB061, BASB063, BASB065, BASB066 and BASB071 polypeptides (see AAY95646 -53) and the polynucleotides (see AAA50070-77) encoding them. These polypeptides and polynucleotides are used to produce medicaments, especially vaccines, to generate immune responses in an animal (claimed).

XX CC Antibodies directed against the polypeptides are used in the treatment of *N. meningitidis* disease (claimed). The polypeptides and polynucleotides can also be used as diagnostic reagents for the diagnosis or staging of *N. meningitidis* disease, and for the detection of an infectious organism response to a drug

XX SQ Sequence 268 AA;

Query Match 81.8%; Score 18; DB 3; Length 268;
Best Local Similarity 57.1%; Pred. No. 3.6e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
| |||
Db 41 VASNAEF 47

RESULT 84
AAU72995
ID AAU72995 standard; protein; 268 AA.

XX AC AAU72995;

XX DT 12-MAR-2002 (first entry)

XX DE *Neisseria meningitidis* virulence protein #85.

XX KW Meningitis; virulence; gene; antibacterial; vaccine; veterinary; infection; Gram-negative bacteria; antimicrobial.

XX OS *Neisseria meningitidis*.

XX PN WO200185772-A2.

XX PD 15-NOV-2001.

XX PF 08-MAY-2001; 2001WO-GB002003.

XX PR 08-MAY-2000; 2000GB-00011108.

XX PA (MICR-) MICROSCIENCE LTD.

XX PI Tang C;

XX DR WPI; 2002-066593/09.

XX DR N-PSDB; AAS97280.

XX PT New peptide encoded by operon including virulence genes of *Neisseria meningitidis*, useful as vaccine component for treating or preventing meningitis and for identifying antimicrobial drug.

XX PS Claim 4; Page 345-346; 423pp; English.

XX CC The invention relates to a peptide (I) encoded by an operon (II) of *Neisseria meningitidis* including virulence genes, or a related molecule having a 40% sequence similarity at the peptide or nucleotide level in a Gram-negative bacterium, or its functional fragment, for therapeutic or diagnostic use. (I) and (II) are useful in the manufacture of a medicament for treating or preventing a condition (e.g., meningitis) associated with infection by *Neisseria* or Gram-negative bacteria. The product is useful for veterinary treatment and in a screening assay for

CC the identification of an antimicrobial drug. The vaccines have prophylactic applications. AAU72911-AAU73014 represent *N. meningitidis* virulence proteins of the invention

XX SQ Sequence 268 AA;

Query Match 81.8%; Score 18; DB 5; Length 268;
Best Local Similarity 57.1%; Pred. No. 3.6e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
| |||
Db 41 VASNAEF 47

RESULT 85
ABU37163
ID ABU37163 standard; protein; 268 AA.

XX AC ABU37163;

XX DT 23-OCT-2003 (revised)

XX DT 19-JUN-2003 (first entry)

XX DE Protein encoded by Prokaryotic essential gene #22690.

XX KW Antisense; prokaryotic essential gene; cell proliferation; drug design.

XX OS *Neisseria gonorrhoeae*.

XX PN WO200277183-A2.

XX PD 03-OCT-2002.

XX PF 21-MAR-2002; 2002WO-US009107.

XX PR 21-MAR-2001; 2001US-00815242.

XX PR 06-SEP-2001; 2001US-00948993.

XX PR 25-OCT-2001; 2001US-0342923P.

XX PR 08-FEB-2002; 2002US-00072851.

XX PR 06-MAR-2002; 2002US-0362699P.

XX PA (ELIT-) ELITRA PHARM INC.

XX PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW; Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX PD WPI; 2003-029926/02.

XX DR N-PSDB; ACA41033.

XX PT New antisense nucleic acids, useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs.

XX PS Claim 25; SEQ ID NO 65087; 1766pp; English.

XX CC The invention relates to an isolated nucleic acid comprising any one of the 6213 antisense sequences given in the specification where expression of the nucleic acid inhibits proliferation of a cell. Also included are: (1) a vector comprising a promoter operably linked to the nucleic acid encoding a polypeptide whose expression is inhibited by the antisense nucleic acid; (2) a host cell containing the vector; (3) an isolated polypeptide or its fragment whose expression is inhibited by the antisense nucleic acid; (4) an antibody capable of specifically binding the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular proliferation; (7) identifying a compound that influences the activity of the gene product or that has an activity against a biological pathway required for proliferation, or that inhibits cellular proliferation; (8) identifying a gene required for cellular proliferation or the biological pathway in which a proliferation-required gene or its gene product lies or a gene on which the test compound that inhibits proliferation of an organism acts; (9) manufacturing an antibiotic; (10) profiling a

CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences. (Updated on 23-OCT-2003 to
 CC standardise OS field)

XX Sequence 268 AA;

SQ Query Match 81.8%; Score 18; DB 6; Length 268;

Best Local Similarity 57.1%; Pred. No. 3.6e+03;

Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXXXAEF 7

Db | | | |

41 VASNAEF 47

RESULT 86

ABU38096

ID ABU38096 standard; protein; 268 AA.

XX AC ABU38096;

XX DT 19-JUN-2003 (first entry)

XX DE Protein encoded by Prokaryotic essential gene #23623.

XX KW Antisense; prokaryotic essential gene; cell proliferation; drug design.

XX OS *Neisseria meningitidis*.

XX PN WO20027183-A2.

XX PD 03-OCT-2002.

XX PF 21-MAR-2002; 2002WO-US009107.

XX PR 21-MAR-2001; 2001US-00815242.

XX PR 06-SEP-2001; 2001US-00948993.

XX PR 25-OCT-2001; 2001US-0342923P.

XX PR 08-FEB-2002; 2002US-00072851.

XX PR 06-MAR-2002; 2002US-0362699P.

XX PA (BLIT-) ELITRA PHARM INC.

XX PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;

XX PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX DR WPI; 2003-0299926/02.

XX DR N-PSDB; ACN41966.

XX XX New antisense nucleic acids, useful for identifying proteins or screening

PT for homologous nucleic acids required for cellular proliferation to

PT isolate candidate molecules for rational drug discovery programs.

XX PS Claim 25; SEQ ID NO 66020; 1766pp; English.

XX CC The invention relates to an isolated nucleic acid comprising any one of

CC the 6213 antisense sequences given in the specification where expression

CC of the nucleic acid inhibits proliferation of a cell. Also included are:

CC (1) a vector comprising a promoter operably linked to the nucleic acid

CC encoding a polypeptide whose expression is inhibited by the antisense

CC nucleic acid; (2) a host cell containing the vector; (3) an isolated

CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC on a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

SQ Sequence 268 AA;

Query Match 81.8%; Score 18; DB 6; Length 268;

Best Local Similarity 57.1%; Pred. No. 3.6e+03;

Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXXXAEF 7

Db | | | |

41 VASNAEF 47

RESULT 87

ABP80010

ID ABP80010 standard; protein; 269 AA.

XX AC ABP80010;

XX DT 07-MAR-2003 (first entry)

XX DE *N. gonorrhoeae* amino acid sequence SEQ ID 6550.

XX KW Antibacterial; infection; vaccine; gene therapy.

XX OS *Neisseria gonorrhoeae*.

XX PN WO200279243-A2.

XX PD 10-OCT-2002.

XX PF 12-FEB-2002; 2002WO-IB002069.

XX PR 12-FEB-2001; 2001GB-00003424.

XX PA (CHIR-) CHIRON SPA.

XX PI Fontana MR, Pizza M, Massignani V, Monaci B;

XX DR WPI; 2003-058415/05.

XX DR N-PSDB; ABZ40980.

XX PT New protein from *Neisseria gonorrhoeae*, useful for the manufacture of a

PT medicament for treating or preventing *N. gonorrhoeae* infection.

XX PS Disclosure; Page 664; 815pp; English.

XX CC The present invention relates to proteins from *Neisseria gonorrhoeae*.

CC Also disclosed are the nucleic acid molecules encoding the proteins and

CC antibodies that specifically bind to the proteins. The composition
 CC comprising the protein, nucleic acid or antibody is useful for the
 CC manufacture of a medicament for treating or preventing N. gonorrhoeae
 CC infection, this may be in the form of a vaccine or gene therapy.
 CC Sequences given in records ABP76736-ABP81046 represent nucleic acid
 CC molecules of the invention
 XX
 XX Sequence 269 AA;
 SQ
 Query Match 81.8%; Score 18; DB 6; Length 269;
 Best Local Similarity 57.1%; Pred. No. 3.6e+03;
 Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 Qy 1 VXXAAEF 7
 Db 42 VASNAEF 48
 RESULT 88
 ADN61450
 ID ADN61450 standard; protein; 269 AA.
 XX
 AC ADN61450;
 XX
 DT 12-AUG-2004 (first entry)
 XX
 DE Human KPP-16 protein SEQ ID NO:16.
 XX
 KW human; kinase; phosphatase; enzyme; KPP; cytostatic;
 KW antiarteriosclerotic; anticonvulsant; nootropic; neuroprotective;
 KW cerebroprotective; anti-HIV; antiallergic; antiinflammatory;
 KW thymomimetic; gene therapy; cell proliferative disorder; cancer;
 KW atherosclerosis; neurological disorder; epilepsy; Huntington's disease;
 KW stroke; immune disorder; inflammatory disorder; AIDS; allergy;
 KW developmental disorder; Hypothyroidism; Cushing's syndrome; infection.
 XX
 OS Homo sapiens.
 XX
 PN WO2004042022-A2.
 XX
 PD 21-MAY-2004.
 XX
 PF 30-OCT-2003; 2003WO-US034809.
 XX
 PR 01-NOV-2002; 2002US-0423226P.
 PR 15-NOV-2002; 2002US-0426713P.
 PR 26-NOV-2002; 2002US-0429766P.
 PR 11-FEB-2003; 2003US-0447043P.
 XX
 PA (INCY-) INCYTE CORP.
 XX
 PI Hafalia AJA, Lee S, Murage J, Swarnakar A, Chawla NK, Khare R;
 PI Elliott VS, Tran UK, Ramkumar J, Gururajan R, Baughn MR, Gietzen KJ;
 PI Yang YG, Chien D, Wang JT, Favero KD, Becha SD, Richardson TW;
 PI Jin P, Hawkins PR, Yue H, Lee EA, Marquis JP;
 XX
 DR WPI; 2004-390608/36.
 DR N-PSDB; ADN61506.
 XX
 XX New human kinases and phosphatases (KPP), useful for diagnosing, treating
 PT and preventing diseases or conditions associated with the aberrant KPP
 PT expression e.g. cancer, AIDS, epilepsy, or infections.
 XX
 XX Claim 1; SEQ ID NO 16; 320pp; English.
 PS
 XX The present sequence represents a human kinase and phosphatase protein
 CC designated KPP-16. Human KPP sequences have cytostatic,
 CC antiarteriosclerotic, anticonvulsant, nootropic, neuroprotective,
 CC cerebroprotective, anti-HIV, antiallergic, antiinflammatory and
 CC thymomimetic activities, and can be used in gene therapy. The human KPP
 CC polypeptides and polynucleotides of the invention are useful in
 CC diagnosing, treating and preventing diseases or conditions associated
 CC with the decreased expression or overexpression of KPP, such as cell

CC proliferative (e.g. cancer, atherosclerosis), neurological (e.g.
 CC epilepsy, Huntington's disease, stroke), immune/inflammatory (e.g. AIDS,
 CC allergies) and developmental (e.g. Hypothyroidism, Cushing's syndrome)
 CC disorders, or infections. They are also useful in assessing the effects
 CC of exogenous compounds on the expression of nucleic acid and amino acid
 CC sequences of KPP. The KPP sequences or their fragments are useful in
 CC screening compounds for effectiveness as agonist or antagonist of the
 CC polypeptides, or in altering the expression of the target polynucleotide
 CC and compounds that specifically bind to or modulate the activity of the
 XX polypeptide.
 XX
 SQ Sequence 269 AA;
 Query Match 81.8%; Score 18; DB 8; Length 269;
 Best Local Similarity 57.1%; Pred. No. 3.6e+03;
 Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 Qy 1 VXXAAEF 7
 Db 228 VITAAEF 234
 RESULT 89
 AAY57069
 ID AAY57069 standard; protein; 271 AA.
 XX
 AC AAY57069;
 XX
 DT 28-FEB-2000 (first entry)
 XX
 DE Calbindin 2 (Calretinin) amino acid sequence.
 XX
 KW Neglected target tissue antigen; NNTA; autoimmunity; autoimmune response;
 KW immunotherapeutic agent; insulin dependent diabetes mellitus;
 KW multiple sclerosis; autoimmune thyroiditis; rheumatoid arthritis;
 KW uveoretinitis; inflammatory response.
 XX
 OS Homo sapiens.
 XX
 PN WO9956763-A1.
 XX
 PD 11-NOV-1999.
 XX
 PF 07-MAY-1999; 99WO-US010250.
 XX
 PR 07-MAY-1998; 98US-0084636P.
 XX
 PA (REGC) UNIV CALIFORNIA.
 XX
 PI Kaufman DL, Tian J, Olcott A;
 XX
 DR WPI; 2000-052905/04.
 XX
 PT Administration of neglected target tissue antigens to modulate immune
 PT responses.
 XX
 PS Disclosure; Page 24; 79pp; English.
 XX
 XX Amino acid sequences AAY57063-Y57091 are examples of neglected target
 CC tissue antigens NNTAs. NNTAs are antigens (whole antigens or fragments)
 CC not involved in autoimmunity. These peptides and proteins are used in the
 CC method of the invention which involves administering an NNTA as an
 CC antigen based immunotherapeutic agent, to a host afflicted with an
 CC autoimmune response associated with an autoimmune disease. The
 CC immunotherapeutic agent is used to treat autoimmune diseases such as
 CC insulin dependent diabetes mellitus, multiple sclerosis, autoimmune
 CC thyroiditis, uveoretinitis, rheumatoid arthritis or abnormal inflammatory
 CC immune responses. The NNTA induces regulatory tolerance by elicitation of
 CC regulatory T cells among T cells recognizing the NNTA but not
 CC participating in the immune response. The NNTA are capable of recognition
 CC by substantial populations of uncommitted T cells which can be primed, or
 CC biased, towards regulatory responses to provide effective treatment. The
 CC NNTA are effective in regulating undesirable immune responses even when

CC target determinants used as agents promoting tolerance agents have failed
 CC to induce an effective regulatory T cell response. NTAS as agents
 CC promoting tolerance are anticipated to be safer than use of target
 CC determinants
 XX
 SQ Sequence 271 AA;

Query Match 81.8%; Score 18; DB 3; Length 271;
 Best Local Similarity 57.1%; Pred. No. 3.7e+03;
 Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VXXAEP 7
 Db 106 VGSSAEP 112

RESULT 90
 AAM78464
 ID AAM78464 standard; protein; 271 AA.
 XX
 AC AAM78464;
 XX
 DT 06-NOV-2001 (first entry)
 XX
 DE Human protein SEQ ID NO 1126.
 XX
 KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorder; arthritis; inflammation.
 XX
 OS Homo sapiens.
 XX
 PN W0200157190-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 05-FEB-2001; 2001WO-US004098.
 XX
 PR 03-FEB-2000; 2000US-00496914.
 PR 27-APR-2000; 2000US-00560875.
 PR 20-JUN-2000; 2000US-00598075.
 PR 19-JUL-2000; 2000US-00620325.
 PR 01-SEP-2000; 2000US-00654936.
 PR 15-SEP-2000; 2000US-00663561.
 PR 20-OCT-2000; 2000US-00693325.
 PR 30-NOV-2000; 2000US-00728422.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;
 PI Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
 PI Xue A, Yang Y, Wejhrman T, Goodrich R;
 XX
 DR WPI; 2001-476283/51.
 DR N-PSDB; AAK51597.
 XX
 PT Nucleic acids encoding polypeptides with cytokine-like activities, useful
 PT in diagnosis and gene therapy.
 XX
 PS Claim 20; Page 3359-3360; 6221pp; English.
 XX
 CC The invention relates to polynucleotides (AAK51456-AAK53435) and the
 CC encoded polypeptides (AAM78333-AAK80302) that exhibit activity elating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activin/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111

CC (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the
 CC sequence listing were missing at the time of publication
 XX
 SQ Sequence 271 AA;

Query Match 81.8%; Score 18; DB 4; Length 271;
 Best Local Similarity 57.1%; Pred. No. 3.7e+03;
 Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VXXAEP 7
 Db 106 VGSSAEP 112

RESULT 91
 ADB70324
 ID ADB70324 standard; protein; 271 AA.
 XX
 AC ADB70324;
 XX
 DT 04-DEC-2003 (first entry)
 XX
 DE Calretinin (calbindin 2) SEQ ID NO:16.
 XX
 KW cancer; malignant pleural mesothelioma; MPM; lung adenocarcinoma;
 KW squamous carcinoma; medulloblastoma; prostate cancer; breast cancer;
 KW diffuse large B-cell lymphoma; follicular lymphoma; ovarian cancer;
 KW human.
 XX
 OS Homo sapiens.
 XX
 PN W02003021229-A2.
 XX
 PD 13-MAR-2003.
 XX
 PF 05-SEP-2002; 2002WO-US028203.
 XX
 PR 05-SEP-2001; 2001US-03117389P.
 PR 30-AUG-2002; 2002US-00236031.
 XX
 PA (BGHM) BRIGHAM & WOMENS HOSPITAL INC.
 XX
 PI Gordon GJ, Jensen RV, Gullans SR, Bueno R;
 DR WPI; 2003-290233/28.
 DR N-PSDB; ADB70323.
 XX
 PT Diagnosing cancer cells in tissue sample, or determining prognosis or
 PT outcome of cancer patient, by calculating ratio of expression levels of
 PT genes that are differentially expressed in cancer and non cancer tissues.
 XX
 PS Claim 77; Page 130-131; 396pp; English.
 XX
 CC The present invention describes a method (M1) for diagnosing the presence
 CC of cancer cells or non-cancer cells in a tissue sample, or determining
 CC the prognosis or outcome of a cancer patient. M1 involves providing a set
 CC of genes that are differentially expressed in cancerous or non-cancerous
 CC conditions, determining the expression levels of the set of genes and
 CC calculating a ratio of the expression levels of the differentially
 CC expressed genes. M1 is useful for diagnosing the presence of cancer cells
 CC or non-cancer cells in a tissue sample, where the cancer is malignant
 CC pleural mesothelioma (MPM), lung adenocarcinoma, squamous carcinoma,
 CC medulloblastoma, prostate cancer, breast cancer, diffuse large B-cell
 CC lymphoma, follicular lymphoma and ovarian cancer, and for determining
 CC prognosis or outcome of a cancer patient. The ratio of expression levels
 CC of differentially expressed genes is used as an indicator of cancer type,
 CC cancer class, and/or cancer prognosis, all of which are useful for
 CC determining a course of treatment of a patient. The present sequence
 CC represents a human protein which is used in an example from the present
 CC invention.
 XX
 SQ Sequence 271 AA;

Query Match 81.8%; Score 18; DB 7; Length 271;
Best Local Similarity 57.1%; Pred. No. 3.7e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
| | | | |
Db 106 VGSSAEF 112

RESULT 93
ADE63979
ID ADE63979 standard; protein; 271 AA.
XX ADE63979;
DE 29-JAN-2004 (first entry)
DE Human Protein P22676, SEQ ID NO 9925.
XX Human; pain; neuronal tissue; gene therapy;
KW spinal segmental nerve injury; chronic constriction injury; CCI;
KW spared nerve injury; SNI; Chung.
XX Homo sapiens.
XX WO2003016475-A2.
XX 27-FEB-2003.
XX 14-AUG-2002; 2002WO-US025765.
XX 14-AUG-2001; 2001US-0312147P.
PR 01-NOV-2001; 2001US-0346382P.
PR 26-NOV-2001; 2001US-0333347P.
XX (GEHO) GEN HOSPITAL CORP.
PA (FARB) BAYER AG.
XX Woolf C, D'urso D, Befort K, Costigan M;
XX WPI; 2003-268312/26.
DR GENBANK; P22676.
XX New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.
XX Claim 1; Page; 1017pp; English.

The invention discloses a composition comprising two or more isolated rat
or human polynucleotides or a polynucleotide which represents a fragment,
derivative or allelic variation of the nucleic acid sequence. Also
claimed are a vector comprising the novel polynucleotide, a host cell
comprising the vector, a method for identifying a nucleotide sequence
which is differentially regulated in an animal subjected to pain and a
kit to perform the method, an array, a method for identifying an agent
that increases or decreases the expression of the polynucleotide sequence
that is differentially expressed in neuronal tissue of a first animal
subjected to pain, a method for identifying a compound which regulates
the expression of a polynucleotide sequence which is differentially
expressed in an animal subjected to pain, a method for identifying a
compound that regulates the activity of one or more of the
polynucleotides, a method for producing a pharmaceutical composition, a
method for identifying a compound or small molecule that regulates the
activity in an animal of one or more of the polypeptides given in the
specification, a method for identifying a compound useful in treating
pain and a pharmaceutical composition comprising the one or more
polypeptides or their antibodies. The polynucleotide or the compound that
modulates its activity is useful for preparing a medicament for treating
pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
therapy). The sequence presented is a human protein (shown in Table 2 of
the specification) which is differentially expressed during pain. Note:
The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic form directly from WIFO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 271 AA;
Query Match 81.8%; Score 18; DB 7; Length 271;
Best Local Similarity 57.1%; Pred. No. 3.7e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
| | | | |
Db 106 VGSSAEF 112

RESULT 93
ADE63983
ID ADE63983 standard; protein; 271 AA.
XX ADE63983;
AC ADE63983;
XX 29-JAN-2004 (first entry)
DT Human Protein P22676, SEQ ID NO 9929.
XX Human; pain; neuronal tissue; gene therapy;
KW spinal segmental nerve injury; chronic constriction injury; CCI;
KW spared nerve injury; SNI; Chung.
XX Homo sapiens.
XX WO2003016475-A2.
XX 27-FEB-2003.
XX 14-AUG-2002; 2002WO-US025765.
XX 14-AUG-2001; 2001US-0312147P.
PR 01-NOV-2001; 2001US-0346382P.
PR 26-NOV-2001; 2001US-0333347P.
XX (GEHO) GEN HOSPITAL CORP.
PA (FARB) BAYER AG.
XX Woolf C, D'urso D, Befort K, Costigan M;
XX WPI; 2003-268312/26.
DR GENBANK; P22676.
XX New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.
XX Claim 1; Page; 1017pp; English.

The invention discloses a composition comprising two or more isolated rat
or human polynucleotides or a polynucleotide which represents a fragment,
derivative or allelic variation of the nucleic acid sequence. Also
claimed are a vector comprising the novel polynucleotide, a host cell
comprising the vector, a method for identifying a nucleotide sequence
which is differentially regulated in an animal subjected to pain and a
kit to perform the method, an array, a method for identifying an agent
that increases or decreases the expression of the polynucleotide sequence
that is differentially expressed in neuronal tissue of a first animal
subjected to pain, a method for identifying a compound which regulates
the expression of a polynucleotide sequence which is differentially
expressed in an animal subjected to pain, a method for identifying a
compound that regulates the activity of one or more of the
polynucleotides, a method for producing a pharmaceutical composition, a
method for identifying a compound or small molecule that regulates the
activity in an animal of one or more of the polypeptides given in the
specification, a method for identifying a compound useful in treating
pain and a pharmaceutical composition comprising the one or more
polypeptides or their antibodies. The polynucleotide or the compound that
modulates its activity is useful for preparing a medicament for treating
pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
therapy). The sequence presented is a human protein (shown in Table 2 of
the specification) which is differentially expressed during pain. Note:
The sequence data for this patent did not form part of the printed

CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a human protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 271 AA;

Query Match 81.8%; Score 18; DB 7; Length 271;
Best Local Similarity 57.1%; Pred. No. 3.7e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
| |||
Db 106 VGSSAEF 112

RESULT 94
ADE63977
ID ADE63977 standard; protein; 271 AA.

XX AC ADE63977;

XX DT 29-JAN-2004 (first entry)

XX DE Rat Protein P47728, SEQ ID NO 9923.

XX KW Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
XX chronic constriction injury; CCI; spared nerve injury; SNI; Chung.

XX OS Rattus norvegicus.

XX PN W02003016475-A2.

XX PD 27-FEB-2003.

XX PF 14-AUG-2002; 2002WO-US025765.

XX PR 14-AUG-2001; 2001US-0312147P.

XX PR 01-NOV-2001; 2001US-0346382P.

XX PR 26-NOV-2001; 2001US-0333347P.

XX PA (GEO) GEN HOSPITAL CORP.

XX PA (FARB) BAYER AG.

XX PI Woolf C, D'urso D, Befort K, Costigan M;

XX WPI; 2003-268312/26.

XX DR GENBANK; P47728.

XX PT New composition comprising two or more isolated polypeptides, useful for
XX preparing a medicament for treating pain in an animal.

XX PS Claim 1; Page; 1017pp; English.

XX CC The invention discloses a composition comprising two or more isolated rat
XX or human polynucleotides or a polynucleotide which represents a fragment,
XX derivative or allelic variation of the nucleic acid sequence. Also
XX claimed are a vector comprising the novel polynucleotide, a host cell
XX comprising the vector, a method for identifying a nucleotide sequence
XX which is differentially regulated in an animal subjected to pain and a
XX kit to perform the method, an array, a method for identifying an agent
XX that increases or decreases the expression of the polynucleotide sequence
XX that is differentially expressed in neuronal tissue of a first animal
XX subjected to pain, a method for identifying a compound which regulates
XX the expression of a polynucleotide sequence which is differentially
XX expressed in an animal subjected to pain, a method for identifying a
XX compound that regulates the activity of one or more of the
XX polynucleotides, a method for producing a pharmaceutical composition, a
XX method for identifying a compound or small molecule that regulates the
XX activity in an animal of one or more of the polypeptides given in the

CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a rat protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 271 AA;

Query Match 81.8%; Score 18; DB 7; Length 271;
Best Local Similarity 57.1%; Pred. No. 3.7e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
| |||
Db 106 VGSSAEF 112

RESULT 95

ADE63981

ID ADE63981 standard; protein; 271 AA.

XX AC ADE63981;

XX DT 29-JAN-2004 (first entry)

XX DE Rat Protein P47728, SEQ ID NO 9927.

XX KW Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
XX chronic constriction injury; CCI; spared nerve injury; SNI; Chung.

XX OS Rattus norvegicus.

XX PN W02003016475-A2.

XX PD 27-FEB-2003.

XX PF 14-AUG-2002; 2002WO-US025765.

XX PR 14-AUG-2001; 2001US-0312147P.

XX PR 01-NOV-2001; 2001US-0346382P.

XX PR 26-NOV-2001; 2001US-0333347P.

XX PA (GEO) GEN HOSPITAL CORP.

XX PA (FARB) BAYER AG.

XX PI Woolf C, D'urso D, Befort K, Costigan M;

XX WPI; 2003-268312/26.

XX DR GENBANK; P47728.

XX PT New composition comprising two or more isolated polypeptides, useful for
XX preparing a medicament for treating pain in an animal.

XX PS Claim 1; Page; 1017pp; English.

XX CC The invention discloses a composition comprising two or more isolated rat
XX or human polynucleotides or a polynucleotide which represents a fragment,
XX derivative or allelic variation of the nucleic acid sequence. Also
XX claimed are a vector comprising the novel polynucleotide, a host cell
XX comprising the vector, a method for identifying a nucleotide sequence
XX which is differentially regulated in an animal subjected to pain and a
XX kit to perform the method, an array, a method for identifying an agent
XX that increases or decreases the expression of the polynucleotide sequence
XX that is differentially expressed in neuronal tissue of a first animal
XX subjected to pain, a method for identifying a compound which regulates
XX the expression of a polynucleotide sequence which is differentially
XX expressed in an animal subjected to pain, a method for identifying a

CC compound that regulates the activity of one or more of the
 CC polynucleotides, a method for producing a pharmaceutical composition, a
 CC method for identifying a compound or small molecule that regulates the
 CC activity in an animal of one or more of the polypeptides given in the
 CC specification, a method for identifying a compound useful in treating
 CC pain and a pharmaceutical composition comprising the one or more
 CC polypeptides or their antibodies. The polynucleotide or the compound that
 CC modulates its activity is useful for preparing a medicament for treating
 CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
 CC therapy). The sequence presented is a rat protein (shown in Table 2 of
 CC the specification) which is differentially expressed during pain. Note:
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic form directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 CC
 SQ Sequence 271 AA;

Query Match 81.8%; Score 18; DB 7; Length 271;
 Best Local Similarity 57.1%; Pred. No. 3.7e+03;
 Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXXAEEF 7
 | |||
 Db 106 VGSSAEF 112

RESULT 96
 ADJ37103
 ID ADJ37103 standard; protein; 271 AA.
 XX
 AC ADJ37103;
 XX
 DT 22-APR-2004 (first entry)
 XX
 DE Human malignant pleural mesothelioma (MPM) protein #4.
 XX
 KW Human; malignant pleural mesothelioma; MPM; tumour; lung adenocarcinoma;
 KW squamous carcinoma; medulloblastoma; prostate cancer; breast cancer;
 KW diffuse large B-cell lymphoma; follicular lymphoma; ovarian cancer;
 KW cytostatic.
 XX
 OS Homo sapiens.
 XX
 PN US2003219760-A1.
 XX
 PD 27-NOV-2003.
 XX
 PF 05-SEP-2002; 2002US-00236031.
 XX
 PR 05-SEP-2001; 2001US-0317389P.
 PR 30-AUG-2002; 2002US-0407431P.
 XX
 PA (BGHM) BRIGHAM & WOMENS HOSPITAL INC.
 XX
 PI Gordon GJ, Jensen RV, Gullans SR, Bueno R;
 XX
 DR WPI; 2004-141744/14.
 DR N-PSDB; ADJ37102.
 XX
 PT Diagnosing the presence of cancer or non-cancer cells in tissue sample,
 PT useful for diagnosing malignant pleural mesothelioma comprises
 PT determining ratio of expression level of a set of genes expressed in
 PT cancer tissues.
 XX
 PS Claim 77; SEQ ID NO 16; 53pp; English.
 XX
 CC The invention relates to a method of diagnosing the presence of cancer
 CC cells or non-cancer cells in a tissue sample, determining prognosis or
 CC outcome of a cancer patient, selecting a course of treatment for a
 CC subject having or suspected of having malignant pleural mesothelioma
 CC (MPM) and evaluating treatment of MPM comprising determining the ratio of
 CC the expression level of a set of genes differentially expressed in a

CC cancer tissue. The cancer is chosen from MPM, lung adenocarcinoma,
 CC squamous carcinoma, medulloblastoma, prostate cancer, breast cancer,
 CC diffuse large B-cell lymphoma, follicular lymphoma and ovarian cancer.
 CC The method is useful for diagnosing MPM in a subject suspected of having
 CC MPM which involves obtaining a tissue sample suspected of being cancerous or
 CC from a subject and determining the expression of nucleic acid markers or
 CC its expression products in the tissue sample. This sequence represents a
 CC human MPM protein of the invention. Note: The sequence data for this
 CC patent did not form part of the printed specification but was obtained in
 CC electronic format directly from USPTO at seqdata.uspto.gov/sequence.html.
 XX
 SQ Sequence 271 AA;

Query Match 81.8%; Score 18; DB 8; Length 271;
 Best Local Similarity 57.1%; Pred. No. 3.7e+03;
 Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXXAEEF 7
 | |||
 Db 106 VGSSAEF 112

RESULT 97
 AAG27922
 ID AAG27922 standard; protein; 284 AA.
 XX
 AC AAG27922;
 XX
 DT 17-OCT-2000 (first entry)
 XX
 DE Arabidopsis thaliana protein fragment SEQ ID NO: 32948.
 XX
 KW Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.
 XX
 OS Arabidopsis thaliana.
 XX
 PN EP1033405-A2.
 XX
 PD 06-SEP-2000.
 XX
 PF 25-FEB-2000; 2000EP-00301439.
 XX
 PR 25-FEB-1999; 99US-0121825P.
 PR 05-MAR-1999; 99US-0123180P.
 PR 09-MAR-1999; 99US-0123548P.
 PR 23-MAR-1999; 99US-0125788P.
 PR 25-MAR-1999; 99US-0126264P.
 PR 29-MAR-1999; 99US-0126785P.
 PR 01-APR-1999; 99US-0127462P.
 PR 06-APR-1999; 99US-0128234P.
 PR 08-APR-1999; 99US-0128714P.
 PR 16-APR-1999; 99US-0129845P.
 PR 19-APR-1999; 99US-0130077P.
 PR 21-APR-1999; 99US-0130449P.
 PR 23-APR-1999; 99US-0130510P.
 PR 28-APR-1999; 99US-0130891P.
 PR 28-APR-1999; 99US-0131449P.
 PR 30-APR-1999; 99US-0132048P.
 PR 30-APR-1999; 99US-0132407P.
 PR 04-MAY-1999; 99US-0132484P.
 PR 05-MAY-1999; 99US-0132485P.
 PR 06-MAY-1999; 99US-0132486P.
 PR 06-MAY-1999; 99US-0132487P.
 PR 07-MAY-1999; 99US-0132863P.
 PR 11-MAY-1999; 99US-0134256P.
 PR 14-MAY-1999; 99US-0134218P.
 PR 14-MAY-1999; 99US-0134219P.
 PR 14-MAY-1999; 99US-0134221P.
 PR 14-MAY-1999; 99US-0134370P.
 PR 18-MAY-1999; 99US-0134768P.
 PR 19-MAY-1999; 99US-0134941P.

PR 20-MAY-1999; 99US-01351124P.
PR 21-MAY-1999; 99US-01353533P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136382P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139452P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139889P.
PR 23-JUN-1999; 99US-0140353P.
PR 24-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140655P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140981P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144332P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145152P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 08-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149829P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 18-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.

CC antigen-presenting cells that express the polypeptide); a method and kit
CC for detecting or determining the presence or absence of P. acnes in a
CC patient; and a method for inhibiting the development of P. acnes in a
CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion
CC proteins, T cell populations or antigen-presenting cells that express the
CC polypeptides are useful for diagnosing, preventing or treating acne
CC vulgaris, or for stimulating an immune response specific for a P. acnes
CC protein. The polynucleotides can also be used as probes or primers for
CC nucleic acid hybridisation. The vaccine composition is useful for the
CC stimulation of an immune response against P. acnes, or for treating acne,
CC and the kit is useful for performing a diagnostic assay. The present
CC sequence represents a specifically claimed P. acnes polypeptide which is
CC thought to contain an immunogenic region. Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 290 AA;

Query Match 81.8%; Score 18; DB 6; Length 290;
Best Local Similarity 57.1%; Pred. No. 3 9e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7

Db 100 VAVTAEF 106

RESULT 100

AAG49548

ID AAG49548 standard; protein; 294 AA.

XX AC AAG49548;

XX DT 18-OCT-2000 (first entry)

DE DE Arabidopsis thaliana protein fragment SEQ ID NO: 62694.

XX KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.

XX OS Arabidopsis thaliana.

XX PN EP1033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-00301439.

XX PR 25-FEB-1999; 99US-0121825P.

XX PR 05-MAR-1999; 99US-0123180P.

XX PR 23-MAR-1999; 99US-0123548P.

XX PR 23-MAR-1999; 99US-0125788P.

XX PR 25-MAR-1999; 99US-0126264P.

XX PR 29-MAR-1999; 99US-0126785P.

XX PR 01-APR-1999; 99US-0127462P.

XX PR 06-APR-1999; 99US-0128234P.

XX PR 08-APR-1999; 99US-0128714P.

XX PR 16-APR-1999; 99US-0129845P.

XX PR 19-APR-1999; 99US-0130077P.

XX PR 21-APR-1999; 99US-0130449P.

XX PR 23-APR-1999; 99US-0130510P.

XX PR 23-APR-1999; 99US-0130891P.

XX PR 28-APR-1999; 99US-0131449P.

XX PR 30-APR-1999; 99US-0132048P.

XX PR 30-APR-1999; 99US-0132407P.

XX PR 04-MAY-1999; 99US-0132484P.

XX PR 05-MAY-1999; 99US-0132485P.

XX PR 06-MAY-1999; 99US-0132486P.

XX PR 07-MAY-1999; 99US-0132487P.

XX PR 11-MAY-1999; 99US-0132863P.

XX PR 11-MAY-1999; 99US-0134256P.

PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139452P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 02-JUL-1999; 99US-0142154P.
PR 06-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.

PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 08-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.

PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 81.8%; Score 18; DB 3; Length 294;
Best Local Similarity 57.1%; Pred. No. 4e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXXAEP 7
| | | | |
Db 72 VTSLAEP 78

Search completed: June 13, 2005, 13:55:56
Job time : 131 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2005, 13:35:12 ; Search time 30 Seconds
(without alignments)
17.418 Million cell updates/sec

Title: 09730329-59ED

Perfect score: 22

Sequence: 1 VXXAAEF 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 300 summaries

Database : Issued Patents AA:*

1: /cgn2_6/ptodata/1/iaa/5A COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCRUS COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	19	86.4	267	3	US-09-302-620B-106
2	19	86.4	267	4	US-09-911-781-24
3	19	86.4	267	4	US-10-400-902-24
4	19	86.4	267	6	5204252-2
5	19	86.4	267	6	5204252-2
6	19	86.4	268	6	5204252-4
7	19	86.4	268	6	5204252-4
8	18	81.8	56	4	US-09-288-143-158
9	18	81.8	69	4	US-09-410-132-4
10	18	81.8	89	4	US-09-949-016-9298
11	18	81.8	179	4	US-09-634-238-296
12	18	81.8	214	4	US-09-489-039A-12873
13	18	81.8	236	3	US-09-605-858-34
14	18	81.8	255	4	US-09-489-039A-10445
15	18	81.8	260	4	US-09-902-540-14188
16	18	81.8	295	4	US-09-252-991A-16774
17	18	81.8	335	4	US-09-583-110-5099
18	18	81.8	338	4	US-09-602-777A-262
19	18	81.8	347	4	US-09-711-164-346
20	18	81.8	360	4	US-09-248-796A-18326
21	18	81.8	386	4	US-09-902-540-15437
22	18	81.8	387	4	US-09-410-132-2
23	18	81.8	397	4	US-09-583-110-4402
24	18	81.8	400	4	US-09-107-433-4600
25	18	81.8	403	4	US-09-489-039A-10042
26	18	81.8	434	2	US-08-529-600D-2
27	18	81.8	434	2	US-08-973-275-4

28	18	81.8	434	3	US-09-122-632-2	Sequence 2, Appli
29	18	81.8	440	4	US-09-684-855-135	Sequence 135, App
30	18	81.8	440	4	US-09-684-855-158	Sequence 158, App
31	18	81.8	440	4	US-09-488-265B-23	Sequence 23, Appl
32	18	81.8	443	4	US-09-328-352-7566	Sequence 7566, Ap
33	18	81.8	475	2	US-08-819-825-2	Sequence 2, Appli
34	18	81.8	475	3	US-09-163-642-2	Sequence 2, Appli
35	18	81.8	475	4	US-09-273-871A-14	Sequence 14, Appl
36	18	81.8	475	4	US-10-083-452-14	Sequence 14, Appl
37	18	81.8	507	4	US-09-949-016-11504	Sequence 11504, A
38	18	81.8	513	2	US-08-459-346-19	Sequence 19, Appl
39	18	81.8	513	2	US-07-989-847-8	Sequence 8, Appli
40	18	81.8	513	3	US-08-889-419-19	Sequence 19, Appl
41	18	81.8	513	3	US-08-469-411-8	Sequence 8, Appli
42	18	81.8	513	3	US-08-402-542-19	Sequence 19, Appl
43	18	81.8	513	4	US-09-780-601A-8	Sequence 8, Appli
44	18	81.8	513	4	US-09-949-016-6118	Sequence 6118, Ap
45	18	81.8	513	5	PCT-US93-07189-19	Sequence 19, Appl
46	18	81.8	513	6	5187076-6	Patent No. 5187076
47	18	81.8	513	6	5187076-6	Patent No. 5187076
48	18	81.8	522	4	US-09-949-016-10103	Sequence 10103, A
49	18	81.8	532	4	US-09-828-302-12	Sequence 12, Appl
50	18	81.8	563	4	US-09-949-016-11259	Sequence 11259, A
51	18	81.8	607	4	US-09-344-882-18	Sequence 18, Appl
52	18	81.8	637	4	US-09-543-681A-5869	Sequence 5869, Ap
53	18	81.8	659	3	US-09-268-347-46	Sequence 46, Appl
54	18	81.8	660	3	US-09-268-347-45	Sequence 45, Appl
55	18	81.8	669	4	US-09-252-991A-31488	Sequence 31488, A
56	18	81.8	672	3	US-09-363-708-2	Sequence 2, Appli
57	18	81.8	672	4	US-09-083-587-2	Sequence 2, Appli
58	18	81.8	711	3	US-08-772-270A-12	Sequence 12, Appl
59	18	81.8	728	4	US-09-569-037-8	Sequence 8, Appli
60	18	81.8	775	1	US-07-603-133B-12	Sequence 12, Appl
61	18	81.8	1010	4	US-09-248-796A-16379	Sequence 16379, A
62	18	81.8	1251	5	PCT-US95-02251-3	Sequence 3, Appli
63	18	81.8	1252	1	US-08-199-780-3	Sequence 3, Appli
64	18	81.8	1252	2	US-08-316-650-3	Sequence 3, Appli
65	18	81.8	1253	3	US-08-479-722B-4	Sequence 4, Appli
66	18	81.8	1253	4	US-09-592-685-4	Sequence 4, Appli
67	18	81.8	2039	4	US-09-077-098A-7	Sequence 7, Appli
68	18	81.8	3913	4	US-09-949-016-10933	Sequence 10933, A
69	18	81.8	4377	4	US-09-949-016-6978	Sequence 6978, Ap
70	17	77.3	7	4	US-09-604-608-27	Sequence 27, Appl
71	17	77.3	8	4	US-09-604-608-28	Sequence 28, Appl
72	17	77.3	9	4	US-09-724-566A-87	Sequence 87, Appl
73	17	77.3	9	4	US-09-724-566A-88	Sequence 88, Appl
74	17	77.3	9	4	US-09-724-566A-89	Sequence 89, Appl
75	17	77.3	9	4	US-09-724-566A-93	Sequence 93, Appl
76	17	77.3	9	4	US-09-724-566A-94	Sequence 94, Appl
77	17	77.3	9	4	US-09-471-669A-87	Sequence 87, Appl
78	17	77.3	9	4	US-09-471-669A-88	Sequence 88, Appl
79	17	77.3	9	4	US-09-471-669A-89	Sequence 89, Appl
80	17	77.3	9	4	US-09-471-669A-93	Sequence 93, Appl
81	17	77.3	9	4	US-09-471-669A-94	Sequence 94, Appl
82	17	77.3	23	4	US-09-634-955B-24	Sequence 24, Appl
83	17	77.3	47	1	US-08-415-751-12	Sequence 12, Appl
84	17	77.3	66	4	US-09-248-796A-24127	Sequence 24127, A
85	17	77.3	68	4	US-09-543-681A-7973	Sequence 7973, Ap
86	17	77.3	68	6	5177197-52	Patent No. 5177197
87	17	77.3	68	6	5177197-52	Patent No. 5177197
88	17	77.3	74	4	US-09-248-796A-16636	Sequence 16636, A
89	17	77.3	78	4	US-08-311-731A-79	Sequence 79, Appl
90	17	77.3	95	2	US-08-479-078-27	Sequence 27, Appl
91	17	77.3	96	2	US-08-479-078-27	Sequence 27, Appl
92	17	77.3	112	4	US-08-311-731A-185	Sequence 185, App
93	17	77.3	122	4	US-09-107-532A-5330	Sequence 5330, Ap
94	17	77.3	136	3	US-09-183-861-49	Sequence 49, Appl
95	17	77.3	136	3	US-09-022-765-49	Sequence 49, Appl
96	17	77.3	136	4	US-09-551-974A-49	Sequence 49, Appl
97	17	77.3	136	4	US-09-565-501A-49	Sequence 49, Appl
98	17	77.3	136	4	US-09-639-206A-49	Sequence 49, Appl
99	17	77.3	143	4	US-09-874-923-49	Sequence 49, Appl
100	17	77.3	143	4	US-09-134-000C-4544	Sequence 4544, Ap
	17	77.3	143	4	US-09-248-796A-14343	Sequence 14343, A

101	17	77.3	161	4	US-09-270-767-43252	Sequence 43252, A	174	17	77.3	371	4	US-09-252-991A-20311	Sequence 20311, A
102	17	77.3	169	4	US-09-248-796A-15773	Sequence 15773, A	175	17	77.3	375	4	US-09-252-991A-24216	Sequence 24216, A
103	17	77.3	182	4	US-09-328-952-4431	Sequence 4431, Ap	176	17	77.3	383	4	US-09-252-991A-25535	Sequence 25535, A
104	17	77.3	185	1	US-07-741-340-6	Sequence 6, Appli	177	17	77.3	385	4	US-09-489-039A-12697	Sequence 12697, A
105	17	77.3	185	1	US-08-289-548A-6	Sequence 6, Appli	178	17	77.3	390	4	US-09-252-991A-21560	Sequence 21560, A
106	17	77.3	185	1	US-08-452-654-6	Sequence 6, Appli	179	17	77.3	392	4	US-09-107-433-3298	Sequence 3298, Ap
107	17	77.3	185	1	US-08-452-655B-6	Sequence 6, Appli	180	17	77.3	421	4	US-09-252-991A-23130	Sequence 23130, Ap
108	17	77.3	185	2	US-08-865-336-3	Sequence 3, Appli	181	17	77.3	424	4	US-09-489-039A-14030	Sequence 12030, A
109	17	77.3	185	2	US-08-865-336-4	Sequence 4, Appli	182	17	77.3	427	2	US-08-846-021A-8	Sequence 8, Appli
110	17	77.3	185	3	US-08-450-582-6	Sequence 6, Appli	183	17	77.3	427	4	US-09-897-425-8	Sequence 8, Appli
111	17	77.3	185	4	US-08-449-731-6	Sequence 6, Appli	184	17	77.3	427	4	US-09-893-525-8	Sequence 8, Appli
112	17	77.3	187	1	US-07-956-700B-8	Sequence 8, Appli	185	17	77.3	432	2	US-08-644-034A-2	Sequence 2, Appli
113	17	77.3	187	1	US-08-476-537-8	Sequence 8, Appli	186	17	77.3	439	4	US-09-328-352-4576	Sequence 4576, Ap
114	17	77.3	187	1	US-08-485-607-8	Sequence 8, Appli	187	17	77.3	440	4	US-09-252-991A-22487	Sequence 22487, Ap
115	17	77.3	187	2	US-08-475-879-8	Sequence 8, Appli	188	17	77.3	446	4	US-09-583-110-5292	Sequence 5292, Ap
116	17	77.3	187	3	US-09-433-043B-8	Sequence 8, Appli	189	17	77.3	452	4	US-09-252-991A-26294	Sequence 26294, A
117	17	77.3	187	3	US-09-248-796A-19854	Sequence 19854, A	190	17	77.3	458	6	5217891-15	Patent No. 5217891
118	17	77.3	189	3	US-08-779-764A-29	Sequence 29, Appl	191	17	77.3	458	6	5217891-15	Patent No. 5217891
119	17	77.3	189	4	US-09-563-456-29	Sequence 29, Appl	192	17	77.3	461	4	US-09-489-039A-10092	Sequence 10092, A
120	17	77.3	194	4	US-09-134-000C-3800	Sequence 3800, Ap	193	17	77.3	462	3	US-09-036-987A-18	Sequence 18, Appl
121	17	77.3	199	4	US-09-328-352-6184	Sequence 6184, Ap	194	17	77.3	462	3	US-09-370-700-18	Sequence 18, Appl
122	17	77.3	206	4	US-09-270-767-58211	Sequence 58211, A	195	17	77.3	462	4	US-09-603-207-18	Sequence 18, Appl
123	17	77.3	210	1	US-07-741-940-4	Sequence 4, Appli	196	17	77.3	462	4	US-09-252-991A-21817	Sequence 21817, A
124	17	77.3	210	1	US-08-289-548A-4	Sequence 4, Appli	197	17	77.3	463	4	US-08-311-731A-7	Sequence 7, Appli
125	17	77.3	210	1	US-08-452-654-4	Sequence 4, Appli	198	17	77.3	469	4	US-09-543-681A-5423	Sequence 5423, Ap
126	17	77.3	210	1	US-08-452-655B-4	Sequence 4, Appli	199	17	77.3	476	4	US-09-248-796A-19192	Sequence 19192, A
127	17	77.3	210	3	US-08-450-582-4	Sequence 4, Appli	200	17	77.3	483	4	US-09-252-991A-17808	Sequence 17808, A
128	17	77.3	210	4	US-08-449-731-4	Sequence 4, Appli	201	17	77.3	487	4	US-09-902-540-11900	Sequence 11900, A
129	17	77.3	212	4	US-09-543-681A-6648	Sequence 6648, Ap	202	17	77.3	490	4	US-09-489-039A-13226	Sequence 13226, A
130	17	77.3	222	1	US-07-956-700B-21	Sequence 21, Appl	203	17	77.3	494	4	US-09-489-039A-10060	Sequence 10060, A
131	17	77.3	222	1	US-08-476-537-21	Sequence 21, Appl	204	17	77.3	497	4	US-09-902-540-10388	Sequence 10388, A
132	17	77.3	222	1	US-08-485-607-21	Sequence 21, Appl	205	17	77.3	504	4	US-09-252-991A-18609	Sequence 18609, A
133	17	77.3	222	2	US-08-475-879-21	Sequence 21, Appl	206	17	77.3	512	3	US-08-856-253-6	Sequence 6, Appli
134	17	77.3	223	3	US-09-433-043B-21	Sequence 21, Appl	207	17	77.3	516	4	US-09-198-452A-598	Sequence 598, App
135	17	77.3	223	4	US-09-489-039A-13347	Sequence 13347, A	208	17	77.3	523	4	US-09-438-185A-562	Sequence 562, App
136	17	77.3	230	4	US-09-107-532A-4412	Sequence 4412, Ap	209	17	77.3	543	4	US-09-252-991A-27650	Sequence 27650, A
137	17	77.3	231	4	US-09-134-000C-3889	Sequence 3889, Ap	210	17	77.3	545	4	US-09-572-147-2	Sequence 2, Appli
138	17	77.3	240	4	US-09-302-626B-170	Sequence 170, App	211	17	77.3	548	4	US-09-107-532A-6627	Sequence 6627, Ap
139	17	77.3	246	4	US-09-540-236-2107	Sequence 2107, Ap	212	17	77.3	551	4	US-09-804-060-2	Sequence 2, Appli
140	17	77.3	248	4	US-09-265-585C-126	Sequence 126, App	213	17	77.3	552	4	US-09-801-191A-2	Sequence 2, Appli
141	17	77.3	250	4	US-09-270-767-38320	Sequence 38320, A	214	17	77.3	552	4	US-09-801-191A-6	Sequence 6, Appli
142	17	77.3	250	4	US-09-270-767-35337	Sequence 35337, A	215	17	77.3	552	4	US-10-345-198-2	Sequence 2, Appli
143	17	77.3	261	4	US-09-328-352-4725	Sequence 4725, Ap	216	17	77.3	552	4	US-10-345-198-6	Sequence 6, Appli
144	17	77.3	267	4	US-09-107-433-3575	Sequence 3575, Ap	217	17	77.3	553	4	US-09-252-991A-19864	Sequence 19864, A
145	17	77.3	273	4	US-09-252-991A-22212	Sequence 22212, A	218	17	77.3	560	4	US-09-134-000C-5182	Sequence 5182, A
146	17	77.3	273	4	US-09-270-767-44034	Sequence 44034, A	219	17	77.3	576	4	US-09-801-191A-4	Sequence 4, Appli
147	17	77.3	283	4	US-09-461-920F-4	Sequence 4, Appli	220	17	77.3	576	4	US-09-801-191A-5	Sequence 5, Appli
148	17	77.3	283	4	US-10-255-018B-4	Sequence 4, Appli	221	17	77.3	576	4	US-10-345-198-4	Sequence 4, Appli
149	17	77.3	285	4	US-09-252-991A-30991	Sequence 30991, A	222	17	77.3	576	4	US-10-345-198-5	Sequence 5, Appli
150	17	77.3	286	3	US-08-588-207A-331	Sequence 331, App	223	17	77.3	579	4	US-09-543-681A-4869	Sequence 4869, Ap
151	17	77.3	286	4	US-09-252-991A-18196	Sequence 18196, A	224	17	77.3	604	4	US-09-328-352-5187	Sequence 5187, Ap
152	17	77.3	287	4	US-09-583-110-4372	Sequence 4372, Ap	225	17	77.3	620	4	US-08-637-670-40	Sequence 40, Appl
153	17	77.3	290	4	US-09-902-540-10278	Sequence 10278, A	226	17	77.3	631	4	US-09-328-352-6860	Sequence 6860, Ap
154	17	77.3	302	4	US-09-543-681A-82770	Sequence 82770, Ap	227	17	77.3	631	4	US-09-538-092-1357	Sequence 1357, Ap
155	17	77.3	313	4	US-09-583-110-4236	Sequence 4236, Ap	228	17	77.3	653	4	US-09-252-991A-27932	Sequence 27932, A
156	17	77.3	315	4	US-09-489-039A-9274	Sequence 9274, Ap	229	17	77.3	661	4	US-09-949-016-7254	Sequence 7254, Ap
157	17	77.3	315	4	US-09-107-433-3706	Sequence 3706, Ap	230	17	77.3	662	3	US-09-405-728-3	Sequence 3, Appli
158	17	77.3	316	4	US-09-134-000C-4732	Sequence 4732, Ap	231	17	77.3	666	4	US-09-270-767-46736	Sequence 46736, A
159	17	77.3	319	4	US-09-107-532A-5442	Sequence 5442, Ap	232	17	77.3	678	4	US-09-270-767-44599	Sequence 44599, A
160	17	77.3	334	4	US-09-710-279-10	Sequence 10, Appl	233	17	77.3	680	4	US-09-252-991A-32736	Sequence 32736, A
161	17	77.3	338	4	US-09-543-681A-5499	Sequence 5499, Ap	234	17	77.3	690	4	US-09-302-626B-171	Sequence 171, App
162	17	77.3	339	3	US-09-134-001C-3913	Sequence 3913, Ap	235	17	77.3	706	1	US-08-484-105-16	Sequence 16, Appl
163	17	77.3	341	3	US-09-134-001C-3568	Sequence 3568, Ap	236	17	77.3	706	1	US-08-484-106-16	Sequence 16, Appl
164	17	77.3	344	1	US-08-403-866-9	Sequence 9, Appli	237	17	77.3	710	4	US-09-302-626B-28	Sequence 28, Appl
165	17	77.3	344	4	US-09-270-767-44219	Sequence 44219, A	238	17	77.3	710	4	US-09-302-626B-30	Sequence 30, Appl
166	17	77.3	352	4	US-09-270-767-40279	Sequence 40279, A	239	17	77.3	713	4	US-09-949-016-10902	Sequence 10902, A
167	17	77.3	352	4	US-09-270-767-55495	Sequence 55495, A	240	17	77.3	719	3	US-08-975-762-59	Sequence 59, Appl
168	17	77.3	356	4	US-09-328-352-7671	Sequence 7671, Ap	241	17	77.3	719	3	US-09-295-028-59	Sequence 59, Appl
169	17	77.3	358	4	US-09-489-039A-14539	Sequence 14539, A	242	17	77.3	719	3	US-09-106-582-59	Sequence 59, Appl
170	17	77.3	360	4	US-09-489-039A-13055	Sequence 13055, A	243	17	77.3	719	4	US-09-159-469-59	Sequence 59, Appl
171	17	77.3	360	4	US-09-583-110-3285	Sequence 3285, Ap	244	17	77.3	719	4	US-09-693-542-59	Sequence 59, Appl
172	17	77.3	360	4	US-09-248-796A-18247	Sequence 18247, A	245	17	77.3	740	4	US-09-252-991A-30687	Sequence 30687, A
173	17	77.3	361	4	US-09-107-433-3390	Sequence 3390, Ap	246	17	77.3	779	4	US-09-252-991A-30151	Sequence 30151, A

247 17 77.3 844 1 US-07-646-537B-2
248 17 77.3 845 4 US-09-538-092-932
249 17 77.3 852 3 US-09-433-043B-118
250 17 77.3 853 3 US-09-433-043B-119
251 17 77.3 860 4 US-09-949-016-7553
252 17 77.3 876 4 US-09-252-991A-32641
253 17 77.3 878 3 US-08-941-936-2
254 17 77.3 898 4 US-09-328-352-6123
255 17 77.3 957 4 US-09-252-991A-21567
256 17 77.3 972 3 US-08-335-844A-24
257 17 77.3 972 4 US-09-129-366-24
258 17 77.3 1183 2 US-08-447-031A-2
259 17 77.3 1394 4 US-09-949-016-5971
260 17 77.3 1394 6 5177197-30
261 17 77.3 1394 6 5177197-30
262 17 77.3 1456 1 US-08-803-973-2
263 17 77.3 1456 1 US-08-803-972-2
264 17 77.3 1651 3 US-09-540-245A-18
265 17 77.3 2972 3 US-09-579-181-2
266 17 77.3 3118 3 US-09-579-181-1
267 17 77.3 3170 3 US-09-036-987A-4
268 17 77.3 3170 3 US-09-370-700-4
269 17 77.3 3170 4 US-09-603-207-4
270 16 72.7 7 2 US-08-792-553-10
271 16 72.7 7 4 US-09-129-192C-44
272 16 72.7 7 4 US-09-724-566A-78
273 16 72.7 7 4 US-08-594-575C-10
274 16 72.7 7 4 US-09-471-669A-78
275 16 72.7 8 4 US-09-548-372D-67
276 16 72.7 8 4 US-09-548-367D-67
277 16 72.7 8 4 US-09-551-853D-67
278 16 72.7 8 4 US-09-724-566A-81
279 16 72.7 8 4 US-09-416-901B-67
280 16 72.7 8 4 US-09-548-376D-67
281 16 72.7 8 4 US-09-794-327A-67
282 16 72.7 8 4 US-09-548-373D-67
283 16 72.7 8 4 US-09-795-847B-67
284 16 72.7 8 4 US-09-869-414-67
285 16 72.7 8 4 US-09-548-366F-67
286 16 72.7 8 4 US-09-548-368D-67
287 16 72.7 8 4 US-09-794-325A-67
288 16 72.7 8 4 US-09-471-669A-81
289 16 72.7 9 3 US-08-802-981-219
290 16 72.7 9 3 US-08-802-981-220
291 16 72.7 9 3 US-08-802-981-221
292 16 72.7 9 3 US-09-294-987-6
293 16 72.7 9 4 US-09-724-566A-52
294 16 72.7 9 4 US-09-724-566A-73
295 16 72.7 9 4 US-09-724-566A-82
296 16 72.7 9 4 US-09-724-566A-83
297 16 72.7 9 4 US-09-724-566A-84
298 16 72.7 9 4 US-09-724-566A-85
299 16 72.7 9 4 US-09-724-566A-86
300 16 72.7 9 4 US-09-724-566A-90

ALIGNMENTS

RESULT 1
US-09-302-620B-106
; Sequence 106, Application US/09302620B
; Patent No. 631420
; GENERAL INFORMATION:
; APPLICANT: Wilson, C. Ron
; APPLICANT: Craft, David L.
; APPLICANT: Eirich, Dudley
; APPLICANT: Eshoo, Mark
; APPLICANT: Madduri, Krishna M.
; APPLICANT: Cornett, Cathy A.
; APPLICANT: Brenner, Alfred A.
; APPLICANT: Tang, Maria
; APPLICANT: Loper, John C.

; APPLICANT: Gleeson, Martin
; TITLE OF INVENTION: CYTOCHROME P450 MONOOXYGENASE AND NADPH CYTOCHROME P450
; TITLE OF INVENTION: OXIDOREDUCTASE GENES AND PROTEINS RELATED TO THE OMEGA
; TITLE OF INVENTION: HYDROXYLASE COMPLEX OF CANDIDA TROPICALIS AND METHODS
; TITLE OF INVENTION: RELATING THERETO
; FILE REFERENCE: 1010-16.seq
; CURRENT APPLICATION NUMBER: US/09/302,620B
; CURRENT FILING DATE: 1999-04-30
; NUMBER OF SEQ ID NOS: 109
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 106
; LENGTH: 267
; TYPE: PRT
; ORGANISM: Candida tropicalis
US-09-302-620B-106

Query Match 86.4%; Score 19; DB 3; Length 267;
Best Local Similarity 57.1%; Pred. No. 4.7e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXXAEP 7
| | | | |
Db 39 VTTTAEF 45

RESULT 2
US-09-911-781-24
; Sequence 24, Application US/09911781
; Patent No. 6673613
; GENERAL INFORMATION:
; APPLICANT: Craft, David L.
; APPLICANT: Wilson, C. Ron
; APPLICANT: Eirich, Dudley
; APPLICANT: Zhang, Yeyan

; TITLE OF INVENTION: USE OF CYP52A2A PROMOTER TO INCREASE GENE EXPRESSION IN YEAST
; FILE REFERENCE: U0012 OS/OAAP (1010-49)
; CURRENT APPLICATION NUMBER: US/09/911,781
; CURRENT FILING DATE: 2001-07-24
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 24
; LENGTH: 267
; TYPE: PRT
; ORGANISM: Candida tropicalis
US-09-911-781-24

Query Match 86.4%; Score 19; DB 4; Length 267;
Best Local Similarity 57.1%; Pred. No. 4.7e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXXAEP 7
| | | | |
Db 39 VTTTAEF 45

RESULT 3
US-10-400-902-24
; Sequence 24, Application US/10400902
; Patent No. 6790640
; GENERAL INFORMATION:
; APPLICANT: Craft, David L.
; APPLICANT: Wilson, C. Ron
; APPLICANT: Eirich, Dudley
; APPLICANT: Zhang, Yeyan

; TITLE OF INVENTION: USE OF CYP52A2A PROMOTER TO INCREASE GENE EXPRESSION IN YEAST
; FILE REFERENCE: U0012 OS/OAAP (1010-49)
; CURRENT APPLICATION NUMBER: US/10/400,902
; CURRENT FILING DATE: 2003-03-31
; PRIOR APPLICATION NUMBER: US/09/911,781
; PRIOR FILING DATE: 2001-07-24
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 24

```
; LENGTH: 267
; TYPE: PRT
; ORGANISM: Candida tropicalis
US-10-400-902-24

Query Match      86.4%; Score 19; DB 4; Length 267;
Best Local Similarity 57.1%; Pred. No. 4.7e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
Db 39 VTTAAEF 45

RESULT 4
5204252-2
; Patent No. 5204252
; APPLICANT: CREGG, JAMES M.; GLEESON, MARTIN A.; HAAS, LISA
; PICATAGGIO, STEPHEN
; TITLE OF INVENTION: CANDIDA TROPICALIS TRANSFORMATION SYSTEM
; NUMBER OF SEQUENCES: 9
; CURRENT APPLICATION NUMBER: US/07/386,837
; FILING DATE: 27-JUL-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 308,481
; FILING DATE: 08-FEB-1989
; SEQ ID NO:2:
; LENGTH: 267
5204252-2

Query Match      86.4%; Score 19; DB 6; Length 267;
Best Local Similarity 57.1%; Pred. No. 4.7e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
Db 39 VTTAAEF 45

RESULT 5
5204252-2
; Patent No. 5204252
; APPLICANT: CREGG, JAMES M.; GLEESON, MARTIN A.; HAAS, LISA
; PICATAGGIO, STEPHEN
; TITLE OF INVENTION: CANDIDA TROPICALIS TRANSFORMATION SYSTEM
; NUMBER OF SEQUENCES: 9
; CURRENT APPLICATION NUMBER: US/07/386,837
; FILING DATE: 27-JUL-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 308,481
; FILING DATE: 08-FEB-1989
; SEQ ID NO:2:
; LENGTH: 267
5204252-2

Query Match      86.4%; Score 19; DB 6; Length 267;
Best Local Similarity 57.1%; Pred. No. 4.7e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
Db 39 VTTAAEF 45

RESULT 6
5204252-4
; Patent No. 5204252
; APPLICANT: CREGG, JAMES M.; GLEESON, MARTIN A.; HAAS, LISA
; PICATAGGIO, STEPHEN
; TITLE OF INVENTION: CANDIDA TROPICALIS TRANSFORMATION SYSTEM
; NUMBER OF SEQUENCES: 9
; CURRENT APPLICATION NUMBER: US/07/386,837
; FILING DATE: 27-JUL-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 308,481
; FILING DATE: 08-FEB-1989
; SEQ ID NO:4:
; LENGTH: 268
5204252-4

Query Match      86.4%; Score 19; DB 6; Length 268;
Best Local Similarity 57.1%; Pred. No. 4.7e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
Db 39 VTTAAEF 45

RESULT 7
5204252-4
; Patent No. 5204252
; APPLICANT: CREGG, JAMES M.; GLEESON, MARTIN A.; HAAS, LISA
; PICATAGGIO, STEPHEN
; TITLE OF INVENTION: CANDIDA TROPICALIS TRANSFORMATION SYSTEM
; NUMBER OF SEQUENCES: 9
; CURRENT APPLICATION NUMBER: US/07/386,837
; FILING DATE: 27-JUL-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 308,481
; FILING DATE: 08-FEB-1989
; SEQ ID NO:4:
; LENGTH: 268
5204252-4

Query Match      86.4%; Score 19; DB 6; Length 268;
Best Local Similarity 57.1%; Pred. No. 4.7e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
Db 39 VTTAAEF 45

RESULT 8
US-09-288-143-158
; Sequence 158, Application US/09288143
; Patent No. 6433139
; GENERAL INFORMATION:
; APPLICANT: Brewer et al.
; TITLE OF INVENTION: 53 Human Secreted Proteins
; FILE REFERENCE: P2018P1
; CURRENT APPLICATION NUMBER: US/09/288,143
; CURRENT FILING DATE: 1999-04-08
; EARLIER APPLICATION NUMBER: PCT/US98/21142
; EARLIER FILING DATE: 1998-10-08
; EARLIER APPLICATION NUMBER: 60/061,463
; EARLIER FILING DATE: 1997-10-09
; EARLIER APPLICATION NUMBER: 60/061,529
; EARLIER FILING DATE: 1997-10-09
; EARLIER APPLICATION NUMBER: 60/071,498
; EARLIER FILING DATE: 1997-10-09
; EARLIER APPLICATION NUMBER: 60/061,527
; EARLIER FILING DATE: 1997-10-09
; EARLIER APPLICATION NUMBER: 60/061,536
; EARLIER FILING DATE: 1997-10-09
; EARLIER APPLICATION NUMBER: 60/061,532
; EARLIER FILING DATE: 1997-10-09
; NUMBER OF SEQ ID NOS: 219
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 158
; LENGTH: 56
```

```

; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (55)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-288-143-158

Query Match      81.8%; Score 18; DB 4; Length 56;
Best Local Similarity 57.1%; Pred. No. 1.7e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VXXAAEF 7
Db 36 VAARAEF 42

RESULT 9
US-09-410-132-4
; Sequence 4, Application US/09410132
; Patent No. 6509458
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel E.
; APPLICANT: Hubert, Rene S.
; APPLICANT: Mitchell, Stephen C.
; TITLE OF INVENTION: NOVEL GENE EXPRESSED IN PROSTATE CANCER
; FILE REFERENCE: 1703-021.US1
; CURRENT APPLICATION NUMBER: US/09/410,132
; CURRENT FILING DATE: 1999-09-30
; EARLIER APPLICATION NUMBER: 60/102,572
; EARLIER FILING DATE: 1998-09-30
; EARLIER APPLICATION NUMBER: 60/146,584
; EARLIER FILING DATE: 1999-07-28
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 69
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-410-132-4

Query Match      81.8%; Score 18; DB 4; Length 69;
Best Local Similarity 42.9%; Pred. No. 2.2e+02;
Matches 3; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VXXAAEF 7
Db 1 ISTTAEF 7

RESULT 10
US-09-949-016-9298
; Sequence 9298, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9298
; LENGTH: 89
; TYPE: PRT
; ORGANISM: Human

```

```
Db      186 VVSTAEF 192
      |||
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 14188
; LENGTH: 260
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
; US-09-902-540-14188

Query Match      81.8%; Score 18; DB 4; Length 260;
Best Local Similarity 57.1%; Pred. No. 9.1e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 VXXAAEF 7
Db      90 VGSAAEF 96
      |||

RESULT 16
US-09-252-991A-16774
; Sequence 16774, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 16774
; LENGTH: 295
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-16774

Query Match      81.8%; Score 18; DB 4; Length 295;
Best Local Similarity 57.1%; Pred. No. 1e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 VXXAAEF 7
Db      240 VVSAAEF 246
      |||

RESULT 17
US-09-583-110-5099
; Sequence 5099, Application US/09583110
; Patent No. 6699703
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
; FILE REFERENCE: PATH00-07A
; CURRENT APPLICATION NUMBER: US/09/583,110
; CURRENT FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/107,433
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/085,131
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: US 60/051,553
; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 5322

Db      186 VVSTAEF 192
      |||
; APPLICANT: Jones, Hal C.
; APPLICANT: Liu, Christopher
; APPLICANT: HULTGREN, Scott J.
; APPLICANT: HRUBY, Dennis E.
; APPLICANT: FRANKS, Christine A.
; APPLICANT: EVANS, Amy K.
; TITLE OF INVENTION: DEGP PERIPLASMIC PROTEASE, A NEW ANTI-INFECTION TARGET
; FILE REFERENCE: 016921-149
; CURRENT APPLICATION NUMBER: US/09/605,858
; CURRENT FILING DATE: 2000-06-29
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 34
; LENGTH: 236
; TYPE: PRT
; ORGANISM: Spneumhtra
; US-09-605-858-34

Query Match      81.8%; Score 18; DB 3; Length 236;
Best Local Similarity 57.1%; Pred. No. 8.2e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 VXXAAEF 7
Db      167 VTTVAEF 173
      |||

RESULT 14
US-09-489-039A-10445
; Sequence 10445, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 10445
; LENGTH: 255
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
; US-09-489-039A-10445

Query Match      81.8%; Score 18; DB 4; Length 255;
Best Local Similarity 57.1%; Pred. No. 9e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 VXXAAEF 7
Db      85 VAQAEF 91
      |||

RESULT 15
US-09-902-540-14188
; Sequence 14188, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
```



```
; SEQ ID NO 5099
; LENGTH: 335
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-583-110-5099

Query Match      81.8%; Score 18; DB 4; Length 335;
Best Local Similarity 57.1%; Pred. No. 1.2e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 VXXAAEF 7
Db      224 VMSAAEF 230

RESULT 18
US-09-602-777A-262
; Sequence 262, Application US/09602777A
; Patent No. 6831165
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Kroger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberhauer, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
; TITLE OF INVENTION: INVOLVED IN HOMEOSTASIS AND ADAPTATION
; FILE REFERENCE: EGI-128CP
; CURRENT APPLICATION NUMBER: US/09/602,777A
; CURRENT FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/141031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: DE 19931636.8
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19932125.6
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932126.4
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932127.2
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932128.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932129.9
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: DE 19932226.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932920.6
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19932922.2
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19932924.9
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19932928.1
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19932930.3
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19932933.8
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19932935.4
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19932973.7
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19933002.6
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19933003.4
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19933005.0
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19933006.9
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19941378.9
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: DE 19941379.7
```

```
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: DE 19941390.8
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: DE 19941391.6
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: DE 19942088.2
; PRIOR FILING DATE: 1999-09-03
; NUMBER OF SEQ ID NOS: 442
; SEQ ID NO 262
; LENGTH: 338
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-602-777A-262

Query Match      81.8%; Score 18; DB 4; Length 338;
Best Local Similarity 57.1%; Pred. No. 1.2e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 VXXAAEF 7
Db      252 VSDTAEF 258

RESULT 19
US-09-711-164-346
; Sequence 346, Application US/09711164
; Patent No. 6589738
; GENERAL INFORMATION:
; APPLICANT: Forsyth, R. Alllyn
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zyskind, Judith
; TITLE OF INVENTION: GENES ESSENTIAL FOR MICROBIAL PROLIFERATION AND ANTISENSE THERAPY
; FILE REFERENCE: ELITRA.008A
; CURRENT APPLICATION NUMBER: US/09/711,164
; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: US 60/164415
; PRIOR FILING DATE: 1999-11-9
; NUMBER OF SEQ ID NOS: 469
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 346
; LENGTH: 347
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-711-164-346

Query Match      81.8%; Score 18; DB 4; Length 347;
Best Local Similarity 42.9%; Pred. No. 1.3e+03;
Matches 3; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY      1 VXXAAEF 7
Db      140 ITSAAEF 146

RESULT 20
US-09-248-796A-18326
; Sequence 18326, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 18326
; LENGTH: 360
; TYPE: PRT
```

; ORGANISM: Candida albicans
US-09-248-796A-18326

Query Match 81.8%; Score 18; DB 4; Length 360;
Best Local Similarity 57.1%; Pred. No. 1.3e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VXXAAEF 7

Db 112 VITSAEF 118

RESULT 21

US-09-902-540-15437
; Sequence 15437, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barty S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 15437
; LENGTH: 386
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-15437

Query Match 81.8%; Score 18; DB 4; Length 386;
Best Local Similarity 57.1%; Pred. No. 1.4e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VXXAAEF 7

Db 218 VAAGAEF 224

RESULT 22

US-09-410-132-2
; Sequence 2, Application US/09410132
; Patent No. 6509458
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel E.
; APPLICANT: Hubert, Rene S.
; APPLICANT: Mitchell, Stephen C.
; TITLE OF INVENTION: NOVEL GENE EXPRESSED IN PROSTATE CANCER
; FILE REFERENCE: 1703-021.US1
; CURRENT APPLICATION NUMBER: US/09/410,132
; CURRENT FILING DATE: 1999-09-30
; EARLIER APPLICATION NUMBER: 60/102,572
; EARLIER FILING DATE: 1998-09-30
; EARLIER APPLICATION NUMBER: 60/146,584
; EARLIER FILING DATE: 1999-07-28
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 387
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-410-132-2

Query Match 81.8%; Score 18; DB 4; Length 387;
Best Local Similarity 42.9%; Pred. No. 1.4e+03;
Matches 3; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VXXAAEF 7

:

Db 211 ISTTAEF 217

RESULT 23

US-09-583-110-4402
; Sequence 4402, Application US/09583110
; Patent No. 6699703
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
; TITLE OF INVENTION: Pneumoniae for Diagnostics and Therapeutics
; FILE REFERENCE: PATH00-07A
; CURRENT APPLICATION NUMBER: US/09/583,110
; CURRENT FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/107,433
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/085,131
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: US 60/051,553
; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 5322
; SEQ ID NO 4402
; LENGTH: 397
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-583-110-4402

Query Match 81.8%; Score 18; DB 4; Length 397;
Best Local Similarity 57.1%; Pred. No. 1.5e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VXXAAEF 7

Db 167 VITVAEF 173

RESULT 24

US-09-107-433-4600
; Sequence 4600, Application US/09107433
; Patent No. 6800744
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID
; SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE
; THERAPEUTICS
; NUMBER OF SEQUENCES: 5206
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: <Unknown>
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: <Unknown>
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,433
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/ 085131
; FILING DATE: May 12, 1998
; APPLICATION NUMBER: 60/051553
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-011
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277

; INFORMATION FOR SEQ ID NO: 4600:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 400 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (B) LOCATION 1...400
; SEQUENCE DESCRIPTION: SEQ ID NO: 4600:
US-09-107-433-4600

Query Match 81.8%; Score 18; DB 4; Length 400;
Best Local Similarity 57.1%; Pred. No. 1.5e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
| | | |
Db 170 VTTVAEF 176

RESULT 25

US-09-489-039A-10042
; Sequence 10042, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:

; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709, 2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 10042
; LENGTH: 403
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-10042

Query Match 81.8%; Score 18; DB 4; Length 403;
Best Local Similarity 57.1%; Pred. No. 1.5e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
| | | |
Db 181 VAESAEP 187

RESULT 26

US-08-529-600D-2
; Sequence 2, Application US/08529600D
; Patent No. 5861285
; GENERAL INFORMATION:

; APPLICANT: Tadashi MATSUNAGA
; TITLE OF INVENTION: PROTEIN-BOUND MAGNETIC PARTICLES AND
; TITLE OF INVENTION: PROCESS OF PRODUCING THE SAME
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PILLSBURY MADISON & SUTRO, L.L.P.
; STREET: 1100 NEW YORK AVENUE, N.W.
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3918
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Microsoft Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/529,600D
; FILING DATE: 18-SEP-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 6-248700
; FILING DATE: 16-SEP-1994
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 434 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-529-600D-2

Query Match 81.8%; Score 18; DB 2; Length 434;
Best Local Similarity 57.1%; Pred. No. 1.6e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
| | | |
Db 57 VATLAEP 63

RESULT 27

US-08-973-275-4
; Sequence 4, Application US/08973275B
; Patent No. 5958706
; GENERAL INFORMATION:

; APPLICANT: MATSUNAGA, Tadashi
; APPLICANT: KAMIYA, Shinji
; APPLICANT: NAMBA, Kenryo
; TITLE OF INVENTION: FINE MAGNETIC PARTICLES CONTAINING USEFUL PROTEINS
; FILE REFERENCE: THEREOF
; CURRENT APPLICATION NUMBER: US/08/973,275B
; CURRENT FILING DATE: 1998-02-09
; EARLIER APPLICATION NUMBER: PCT/JP97/01043
; EARLIER FILING DATE: 1997-03-27
; EARLIER APPLICATION NUMBER: JP 8-97536
; EARLIER FILING DATE: 1996-03-28
; EARLIER APPLICATION NUMBER: JP 8-146833
; EARLIER FILING DATE: 1996-05-16
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 4
; LENGTH: 434
; TYPE: PRT
; ORGANISM: Magnetospirillum
; FEATURE:
; NAME/KEY: TRANSMEM
; LOCATION: (7)..(380)
; OTHER INFORMATION: Membrane bound region of mag A protein.
US-08-973-275-4

Query Match 81.8%; Score 18; DB 2; Length 434;
Best Local Similarity 57.1%; Pred. No. 1.6e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
| | | |
Db 57 VATLAEP 63

RESULT 28

US-09-122-632-2
; Sequence 2, Application US/09122632
; Patent No. 6033878
; GENERAL INFORMATION:

; APPLICANT: Tadashi MATSUNAGA
; TITLE OF INVENTION: PROTEIN-BOUND MAGNETIC PARTICLES AND

```
; TITLE OF INVENTION: PROCESS OF PRODUCING THE SAME
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PILLSBURY MADISON & SUTRO, L.L.P.
; STREET: 1100 NEW YORK AVENUE, N.W.
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3918
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Microsoft Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/122,632
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/529,600
; FILING DATE: 18-SEP-1995
; APPLICATION NUMBER: JP 6-248700
; FILING DATE: 16-SEP-1994
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 434 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-122-632-2
```

```
Query Match 81.8%; Score 18; DB 3; Length 434;
Best Local Similarity 57.1%; Pred. No. 1.6e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
| |||
Db 57 VAVIAEF 63
```

```
RESULT 29
US-09-684-855-135
; Sequence 135, Application US/09684855
; Patent No. 6599735
; GENERAL INFORMATION:
; APPLICANT: F. Hoffmann-La Roche AG
; TITLE OF INVENTION: CONTINUOUS FERMENTATION PROCESS
; FILE REFERENCE: C38435/111692
; CURRENT APPLICATION NUMBER: US/09/684,855
; CURRENT FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: EP 00121663.9
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: EP 99120289.6
; PRIOR FILING DATE: 1999-10-11
; NUMBER OF SEQ ID NOS: 169
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 135
; LENGTH: 440
; TYPE: PRT
; ORGANISM: T. lanuginosa
; US-09-684-855-135
```

```
Query Match 81.8%; Score 18; DB 4; Length 440;
Best Local Similarity 57.1%; Pred. No. 1.6e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
| |||
Db 126 VIAAEF 132
```

```
RESULT 30
US-09-684-855-158
```

```
; Sequence 158, Application US/09684855
; Patent No. 6599735
; GENERAL INFORMATION:
; APPLICANT: F. Hoffmann-La Roche AG
; TITLE OF INVENTION: CONTINUOUS FERMENTATION PROCESS
; FILE REFERENCE: C38435/111692
; CURRENT APPLICATION NUMBER: US/09/684,855
; CURRENT FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: EP 00121663.9
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: EP 99120289.6
; PRIOR FILING DATE: 1999-10-11
; NUMBER OF SEQ ID NOS: 169
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 158
; LENGTH: 440
; TYPE: PRT
; ORGANISM: T. lanuginosa
; US-09-684-855-158
```

```
Query Match 81.8%; Score 18; DB 4; Length 440;
Best Local Similarity 57.1%; Pred. No. 1.6e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
| |||
Db 126 VIAAEF 132
```

```
RESULT 31
US-09-488-265B-23
; Sequence 23, Application US/09488265B
; Patent No. 6720174
; GENERAL INFORMATION:
; APPLICANT: Lehmann, Martin
; APPLICANT: Lassen, Soren F
; TITLE OF INVENTION: Improved Phytases
; FILE REFERENCE: 5808.200-US
; CURRENT APPLICATION NUMBER: US/09/488,265B
; CURRENT FILING DATE: 2000-01-20
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 23
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Thermomyces lanuginosus
; US-09-488-265B-23
```

```
Query Match 81.8%; Score 18; DB 4; Length 440;
Best Local Similarity 57.1%; Pred. No. 1.6e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
| |||
Db 126 VIAAEF 132
```

```
RESULT 32
US-09-328-352-7566
; Sequence 7566, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 7566
; LENGTH: 443
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
```

US-09-328-352-7566

Query Match 81.8%; Score 18; DB 4; Length 443;
Best Local Similarity 57.1%; Pred. No. 1.6e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
| | | | |
Db 351 VAGTAEF 357

RESULT 33

US-08-819-825-2

; Sequence 2, Application US/08819825

; Patent No. 5866118

; GENERAL INFORMATION:

; APPLICANT: Berkka, Randy M.

; APPLICANT: Ray, Michael W.

; APPLICANT: Klotz, Alan V.

; TITLE OF INVENTION: Polypeptides Having Phytase Activity

; TITLE OF INVENTION: And Nucleic Acids Encoding Same

; NUMBER OF SEQUENCES: 5

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: No. 58661180 No. 5866118disk of No. 5866118th America, Inc.

; STREET: 405 Lexington Avenue, Suite 6400

; CITY: New York

; STATE: New York

; COUNTRY: U.S.A.

; ZIP: 10174-6401

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/819,825

; FILING DATE: 18-MAR-1997

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Lambiris, Elias J.

; REGISTRATION NUMBER: 33,728

; REFERENCE/DOCKET NUMBER: 4758.200-US

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 212 867 0123

; TELEFAX: 212 867 0298

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 475 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; FRAGMENT TYPE: internal

US-08-819-825-2

Query Match 81.8%; Score 18; DB 2; Length 475;

Best Local Similarity 57.1%; Pred. No. 1.8e+03;

Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
| | | | |
Db 161 VIASAEF 167

RESULT 34

US-09-163-642-2

; Sequence 2, Application US/09163642

; Patent No. 6221644

; GENERAL INFORMATION:

; APPLICANT: Berkka, Randy M.

; APPLICANT: Ray, Michael W.

; APPLICANT: Klotz, Alan V.

; TITLE OF INVENTION: Polypeptides Having Phytase Activity

; TITLE OF INVENTION: And Nucleic Acids Encoding Same

; NUMBER OF SEQUENCES: 5

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: No. 62216440 No. 6221644disk of No. 6221644th America, Inc.

; STREET: 405 Lexington Avenue, Suite 6400

; CITY: New York

; STATE: New York

; COUNTRY: U.S.A.

; ZIP: 10174-6401

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/163,642

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/08/819,825

; FILING DATE: 18-MAR-1997

; ATTORNEY/AGENT INFORMATION:

; NAME: Lambiris, Elias J.

; REGISTRATION NUMBER: 33,728

; REFERENCE/DOCKET NUMBER: 4758.200-US

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 212 867 0123

; TELEFAX: 212 867 0298

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 475 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; FRAGMENT TYPE: internal

US-09-163-642-2

Query Match 81.8%; Score 18; DB 3; Length 475;

Best Local Similarity 57.1%; Pred. No. 1.8e+03;

Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
| | | | |
Db 161 VIASAEF 167

RESULT 35

US-09-273-871A-14

; Sequence 14, Application US/09273871A

; Patent No. 6514495

; GENERAL INFORMATION:

; APPLICANT: Svendsen, Allan

; TITLE OF INVENTION: Phytase Variants

; FILE REFERENCE: 5618.500-US

; CURRENT APPLICATION NUMBER: US/09/273,871A

; CURRENT FILING DATE: 1999-03-22

; PRIOR APPLICATION NUMBER: PA 1998 00407

; PRIOR FILING DATE: 1998-03-23

; PRIOR APPLICATION NUMBER: PA 1998 00806

; PRIOR FILING DATE: 1998-06-19

; PRIOR APPLICATION NUMBER: PA 1998 01176

; PRIOR FILING DATE: 1998-09-18

; PRIOR APPLICATION NUMBER: PA 1999 00091

; PRIOR FILING DATE: 1999-01-22

; PRIOR APPLICATION NUMBER: 60/080,129

; PRIOR FILING DATE: 1998-03-31

; PRIOR APPLICATION NUMBER: 60/090,675

; PRIOR FILING DATE: 1998-06-25

; NUMBER OF SEQ ID NOS: 15

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 14

; LENGTH: 475

```
; TYPE: PRT
; ORGANISM: Thermomyces lanuginosa
US-09-273-871A-14

Query Match      81.8%; Score 18; DB 4; Length 475;
Best Local Similarity 57.1%; Pred. No. 1.8e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VXXAAEF 7
Db 161 VIASAEF 167

RESULT 36
US-10-083-452-14
; Sequence 14, Application US/10083452
; Patent No. 6689358
; GENERAL INFORMATION:
; APPLICANT: Svendsen, Allan
; TITLE OF INVENTION: Phytase Variants
; FILE REFERENCE: 5618.500-US
; CURRENT APPLICATION NUMBER: US/10/083.452
; CURRENT FILING DATE: 2002-02-26
; PRIOR APPLICATION NUMBER: US/09/273,871
; PRIOR FILING DATE: 1999-03-22
; PRIOR APPLICATION NUMBER: PA 1998 00407
; PRIOR FILING DATE: 1998-03-23
; PRIOR APPLICATION NUMBER: PA 1998 00806
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: PA 1998 01176
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: PA 1999 00091
; PRIOR FILING DATE: 1999-01-22
; PRIOR APPLICATION NUMBER: 60/080,129
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/090,675
; PRIOR FILING DATE: 1998-06-25
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 475
; TYPE: PRT
; ORGANISM: Thermomyces lanuginosa
US-10-083-452-14

Query Match      81.8%; Score 18; DB 4; Length 475;
Best Local Similarity 57.1%; Pred. No. 1.8e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VXXAAEF 7
Db 161 VIASAEF 167

RESULT 37
US-09-949-016-11504
; Sequence 11504, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 11504
; LENGTH: 507
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-11504

Query Match      81.8%; Score 18; DB 4; Length 507;
Best Local Similarity 57.1%; Pred. No. 1.9e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VXXAAEF 7
Db 288 VITAAEF 294

RESULT 38
US-08-459-346-19
; Sequence 19, Application US/08459346
; Patent No. 5834179
; GENERAL INFORMATION:
; APPLICANT: JONES, WILLIAM K
; APPLICANT: TUCKER, RONALD F
; APPLICANT: RUEGER, DAVID C
; APPLICANT: OPPERMANN, HERMANN
; APPLICANT: OZKAYNAK, ENGIN
; APPLICANT: KUBERASAMPATH, THANGAVEL
; TITLE OF INVENTION: NOVEL MORPHOGENIC PROTEIN COMPOSITIONS
; TITLE OF INVENTION: OF MATTER
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PATENT ADMINISTRATOR/CREATIVE BIOMOLECULES,
; ADDRESSEE: INC.
; STREET: 35 SOUTH STREET
; CITY: HOPKINTON
; STATE: MA
; COUNTRY: USA
; ZIP: 01748
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/459,346
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/029,335
; FILING DATE: 04-MAR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/971,091
; FILING DATE: 03-NOV-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/946,235
; FILING DATE: 16-SEP-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/938,336
; FILING DATE: 08-AUG-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/923,780
; FILING DATE: 31-JUL-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: FITCHER, EDMUND R
; REGISTRATION NUMBER: 27,829
; REFERENCE/DOCKET NUMBER: CRP-081CP
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 513 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
```

; NAME/KEY: Protein
; LOCATION: 1..513
; OTHER INFORMATION: /note= "PRE-PRO-BMP6 (HUMAN)"
; PUBLICATION INFORMATION:
; AUTHORS: CELESTE,
; JOURNAL: Proc. Natl. Acad. Sci. U.S.A.
; VOLUME: 87
; PAGES: 9843-9847
; DATE: 1991
; US-08-459-346-19

Query Match 81.8%; Score 18; DB 2; Length 513;
Best Local Similarity 57.1%; Pred. No. 1.9e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
| | | | |
Db 250 VVTADEF 256

RESULT 39
US-07-989-847-8
; Sequence 8, Application US/07989847
; Patent No. 5868364
; GENERAL INFORMATION:
; APPLICANT: Israel, David
; APPLICANT: Wolfman, Neil M.
; TITLE OF INVENTION: Recombinant Bone Morphogenetic Protein
; TITLE OF INVENTION: Heterodimers, Compositions and Methods of Use.
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Legal Affairs, Genetics Institute, Inc.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: MA
; COUNTRY: USA
; ZIP: 02140-2387

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Tape
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/989,847
; FILING DATE:
; CLASSIFICATION: 436
; ATTORNEY/AGENT INFORMATION:
; NAME: Kapinos, Ellen J.
; REGISTRATION NUMBER: 32,245
; REFERENCE/DOCKET NUMBER: GI-5192B
; TELEPHONE: 617-876-1170
; TELEFAX: 617-876-5851
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 513 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-989-847-8

Query Match 81.8%; Score 18; DB 2; Length 513;
Best Local Similarity 57.1%; Pred. No. 1.9e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
| | | | |
Db 250 VVTADEF 256

RESULT 40
US-08-889-419-19
; Sequence 19, Application US/08889419

; Patent No. 6071708
; GENERAL INFORMATION:
; APPLICANT: JONES, WILLIAM K
; APPLICANT: TUCKER, RONALD F
; APPLICANT: RUEGER, DAVID C
; APPLICANT: OPPERMAN, HERMANN
; APPLICANT: OZKAYNAK, ENGIN
; APPLICANT: KUBERASAMPATH, THANGAVEL
; TITLE OF INVENTION: NOVEL MORPHOGENIC PROTEIN COMPOSITIONS
; TITLE OF INVENTION: OF MATTER
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patent Administrator, Testa, Hurwitz &
; ADDRESSEE: Thibeault, LLP
; STREET: 125 HIGH STREET
; CITY: BOSTON
; STATE: MA
; COUNTRY: USA
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/889,419
; FILING DATE: 08-JUL-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/459,346
; FILING DATE: 02-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: MEYERS, THOMAS C.
; REGISTRATION NUMBER: 36,989
; REFERENCE/DOCKET NUMBER: CRP-081DVCN
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 513 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..513
; OTHER INFORMATION: /note= "PRE-PRO-BMP6 (HUMAN)"
; PUBLICATION INFORMATION:
; AUTHORS: CELESTE,
; JOURNAL: Proc. Natl. Acad. Sci. U.S.A.
; VOLUME: 87
; PAGES: 9843-9847
; DATE: 1991
; US-08-889-419-19

Query Match 81.8%; Score 18; DB 3; Length 513;
Best Local Similarity 57.1%; Pred. No. 1.9e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
| | | | |
Db 250 VVTADEF 256

RESULT 41
US-08-469-411-8
; Sequence 8, Application US/08469411
; Patent No. 6130880
; GENERAL INFORMATION:
; APPLICANT: Israel, David
; APPLICANT: Wolfman, Neil M.
; TITLE OF INVENTION: Recombinant Bone Morphogenetic Protein
; Heterodimers, Compositions and Methods of Use.
; NUMBER OF SEQUENCES: 30

;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Legal Affairs, Genetics Institute, Inc.
;; STREET: 87 Cambridgepark Drive
;; CITY: Cambridge
;; STATE: MA
;; COUNTRY: USA
;; ZIP: 02140-2387
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Tape
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patent in Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/469,411
;; FILING DATE: 06-Jun-1995
;; CLASSIFICATION: <Unknown>
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Kapinos, Ellen J.
;; REGISTRATION NUMBER: 32,245
;; REFERENCE/DOCKET NUMBER: GI-5192B-CON
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 617-498-8622
;; TELEFAX: 617-876-5851
;; INFORMATION FOR SEQ ID NO: 8:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 513 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-08-469-411-8

Query Match 81.8%; Score 18; DB 3; Length 513;
Best Local Similarity 57.1%; Pred. No. 1.9e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
| |||
Db 250 VVTAAEF 256

RESULT 42
US-08-402-542-19
; Sequence 19, Application US/08402542
; Patent No. 6395883
; GENERAL INFORMATION:
; APPLICANT: JONES, WILLIAM K
; APPLICANT: TUCKER, RONALD F
; APPLICANT: RUEGER, DAVID C
; APPLICANT: OPPERMANN, HERMANN
; APPLICANT: OZKAYNAK, ENGIN
; APPLICANT: KUBERASAMPATH, THANGAVEL
; TITLE OF INVENTION: NOVEL MORPHOGENIC PROTEIN COMPOSITIONS
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PATENT ADMINISTRATOR/CREATIVE BIOMOLECULES,
; INC.
; STREET: 35 SOUTH STREET
; CITY: HOPKINTON
; STATE: MA
; COUNTRY: USA
; ZIP: 01748
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/402,542
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:

;; APPLICATION NUMBER: US/08/040,510
;; FILING DATE: US 08/029,335
;; APPLICATION NUMBER: US 08/029,335
;; FILING DATE: 04-MAR-1993
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/971,091
;; FILING DATE: 03-NOV-1992
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/946,235
;; FILING DATE: 16-SEP-1992
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/938,336
;; FILING DATE: 08-AUG-1992
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/923,780
;; FILING DATE: 31-JUL-1992
;; ATTORNEY/AGENT INFORMATION:
;; NAME: PITCHER, EDMUND R.
;; REGISTRATION NUMBER: 27,829
;; REFERENCE/DOCKET NUMBER: CRP-081CP
;; INFORMATION FOR SEQ ID NO: 19:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 513 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; FEATURE:
;; NAME/KEY: Protein
;; LOCATION: 1..513
;; OTHER INFORMATION: /note= "PRE-PRO-BMP6 (HUMAN)"
;; PUBLICATION INFORMATION:
;; AUTHORS: CELESTE,
;; JOURNAL: Proc. Natl. Acad. Sci. U.S.A.
;; VOLUME: 87
;; PAGES: 9843-9847
;; DATE: 1991
;; US-08-402-542-19

Query Match 81.8%; Score 18; DB 3; Length 513;
Best Local Similarity 57.1%; Pred. No. 1.9e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
| |||
Db 250 VVTAAEF 256

RESULT 43
US-09-780-601A-8
; Sequence 8, Application US/09780601A
; Patent No. 6593109
; GENERAL INFORMATION:
; APPLICANT: Israel, David
; APPLICANT: Wolfman, Neil M.
; TITLE OF INVENTION: Recombinant Bone Morphogenetic Protein
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Legal Affairs, Genetics Institute, Inc.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: MA
; COUNTRY: USA
; ZIP: 02140-2387
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Tape
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/780,601A
; FILING DATE:


```
;
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/469,411
; FILING DATE: 06-Jun-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Kapinos, Ellen J.
; REGISTRATION NUMBER: 32,245
; REFERENCE/DOCKET NUMBER: GI-5192B-CON
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-498-8622
; TELEFAX: 617-876-5851
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 513 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-780-601A-8

Query Match      81.8%; Score 18; DB 4; Length 513;
Best Local Similarity 57.1%; Pred. No. 1.9e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 VXXAAEF 7
.      | | | |
Db      250 VTAAEF 256

RESULT 44
US-09-949-016-6118
; Sequence 6118, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6118
; LENGTH: 513
; TYPE: PRT
; ORGANISM: Human
; US-09-949-016-6118

Query Match      81.8%; Score 18; DB 4; Length 513;
Best Local Similarity 57.1%; Pred. No. 1.9e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 VXXAAEF 7
.      | | | |
Db      250 VTAAEF 256

RESULT 45
PCT-US93-07189-19
; Sequence 19, Application PC/TUS9307189
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: NOVEL MORPHOGENIC PROTEIN COMPOSITIONS
; TITLE OF INVENTION: OF MATTER
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PATENT ADMINISTRATOR/CREATIVE BIOMOLECULES,
; ADDRESSEE: INC.
```

```
;
; STREET: 35 SOUTH STREET
; CITY: HOPKINTON
; STATE: MA
; COUNTRY: USA
; ZIP: 01748
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/07189
; FILING DATE: 19930729
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: KELLEY, ROBIN D.
; REGISTRATION NUMBER: 34,637
; REFERENCE/DOCKET NUMBER: CRP-081CP
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 513 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..513
; OTHER INFORMATION: /note= "PRE-PRO-BMP6 (HUMAN)"
; PUBLICATION INFORMATION:
; AUTHORS: CELESTE,
; JOURNAL: Proc. Natl. Acad. Sci. U.S.A.
; VOLUME: 87
; PAGES: 9843-9847
; DATE: 1991
; PCT-US93-07189-19

Query Match      81.8%; Score 18; DB 5; Length 513;
Best Local Similarity 57.1%; Pred. No. 1.9e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 VXXAAEF 7
.      | | | |
Db      250 VTAAEF 256

RESULT 46
5187076-6
; Patent No. 5187076
; APPLICANT: WOZNEY, JOHN M.; WANG, ELIZABETH A.; ROSEN, VICKI A.;
; CELESTE, ANTHONY J.
; TITLE OF INVENTION: DNA SEQUENCES ENCODING BMP-6 PROTEINS
; NUMBER OF SEQUENCES: 16
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/490,033
; FILING DATE: 07-MAR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 370,544
; FILING DATE: 23-JUN-1989
; APPLICATION NUMBER: 347,559
; FILING DATE: 04-MAY-1989
; APPLICATION NUMBER: 329,610
; FILING DATE: 28-MAR-1989
; APPLICATION NUMBER: 179,100
; FILING DATE: 08-APR-1988
; APPLICATION NUMBER: 179,101
; FILING DATE: 08-APR-1988
; APPLICATION NUMBER: 179,197
; FILING DATE: 08-APR-1988
; APPLICATION NUMBER: 28,285
```

```
; FILING DATE: 20-MAR-1987
; APPLICATION NUMBER: 31,346
; FILING DATE: 26-MAR-1987
; APPLICATION NUMBER: 943,322
; FILING DATE: 17-DEC-1986
; APPLICATION NUMBER: 880,776
; FILING DATE: 01-JUL-1986
; SEQ ID NO:6
; LENGTH: 513
5187076-6

Query Match      81.8%; Score 18; DB 6; Length 513;
Best Local Similarity 57.1%; Pred. No. 1.9e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy      1 VXXAAEF 7
Db      250 VVTADEF 256

RESULT 47
5187076-6
; Patent No. 5187076
; APPLICANT: WOZNEY, JOHN M.; WANG, ELIZABETH A.; ROSEN, VICKI A.;
; CELESTE, ANTHONY J.
; TITLE OF INVENTION: DNA SEQUENCES ENCODING BMP-6 PROTEINS
; NUMBER OF SEQUENCES: 16
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/490,033
; FILING DATE: 07-MAR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 370,544
; FILING DATE: 23-JUN-1989
; APPLICATION NUMBER: 347,559
; FILING DATE: 04-MAY-1989
; APPLICATION NUMBER: 329,610
; FILING DATE: 28-MAR-1989
; APPLICATION NUMBER: 179,100
; FILING DATE: 08-APR-1988
; APPLICATION NUMBER: 179,101
; FILING DATE: 08-APR-1988
; APPLICATION NUMBER: 179,197
; FILING DATE: 08-APR-1988
; APPLICATION NUMBER: 28,285
; FILING DATE: 20-MAR-1987
; APPLICATION NUMBER: 31,346
; FILING DATE: 26-MAR-1987
; APPLICATION NUMBER: 943,322
; FILING DATE: 17-DEC-1986
; APPLICATION NUMBER: 880,776
; FILING DATE: 01-JUL-1986
; SEQ ID NO:6
; LENGTH: 513
5187076-6

Query Match      81.8%; Score 18; DB 6; Length 513;
Best Local Similarity 57.1%; Pred. No. 1.9e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy      1 VXXAAEF 7
Db      250 VVTADEF 256

RESULT 48
US-09-949-016-10103
; Sequence 10103, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
```

```
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10103
; LENGTH: 522
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-10103

Query Match      81.8%; Score 18; DB 4; Length 522;
Best Local Similarity 57.1%; Pred. No. 2e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy      1 VXXAAEF 7
Db      303 VITAAEF 309

RESULT 49
US-09-828-302-12
; Sequence 12, Application US/09828302
; Patent No. 6818805
; GENERAL INFORMATION:
; APPLICANT: COSTA E SILVA, OSWALDO DA
; APPLICANT: VAN THIELEN, NOCHA
; APPLICANT: CHEN, ROUYING
; APPLICANT: ISHITANI, MANABU
; TITLE OF INVENTION: PHOSPHATASE STRESS-RELATED PROTEINS AND METHODS OF USE
; TITLE OF INVENTION: IN PLANTS
; FILE REFERENCE: 16313-0029
; CURRENT APPLICATION NUMBER: US/09/828,302
; CURRENT FILING DATE: 2001-08-20
; PRIOR APPLICATION NUMBER: 60/196,001
; PRIOR FILING DATE: 2000-04-07
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 532
; TYPE: PRT
; ORGANISM: Physcomitrella patens
US-09-828-302-12

Query Match      81.8%; Score 18; DB 4; Length 532;
Best Local Similarity 57.1%; Pred. No. 2e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy      1 VXXAAEF 7
Db      298 VITSAEF 304

RESULT 50
US-09-949-016-11259
; Sequence 11259, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
```

; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11259
; LENGTH: 563
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-11259

Query Match 81.8%; Score 18; DB 4; Length 563;
Best Local Similarity 57.1%; Pred. No. 2.1e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
|
Db 300 VVTADEF 306

RESULT 51

US-09-344-882-18
; Sequence 18, Application US/09344882
; Patent No. 6764851

; GENERAL INFORMATION:
; APPLICANT: Nikolau, Basil J

; APPLICANT: Wurtelle, Eve S

; APPLICANT: Oliver, David J

; APPLICANT: Behal, Robert

; APPLICANT: Schnable, Patrick S

; APPLICANT: Ke, Jinshan

; APPLICANT: Johnson, Jerry L

; APPLICANT: Allred, Carolyn C

; APPLICANT: Fatland, Beth

; APPLICANT: Latzinger, Isabelle

; APPLICANT: Wen, Tsui-Jung

; TITLE OF INVENTION: Materials and Methods for the Alteration of Enzyme and

; FILE REFERENCE: 201573

; CURRENT APPLICATION NUMBER: US/09/344,882

; CURRENT FILING DATE: 1999-06-25

; PRIOR FILING DATE: 1998-06-26

; NUMBER OF SEQ ID NOS: 38

; SOFTWARE: PatentIn Ver. 2.2

; SEQ ID NO 18

; LENGTH: 607

; TYPE: PRT

; ORGANISM: Arabidopsis Thaliana

US-09-344-882-18

Query Match 81.8%; Score 18; DB 4; Length 607;
Best Local Similarity 57.1%; Pred. No. 2.3e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
|
Db 243 VEADEF 249

RESULT 52

US-09-543-681A-5869

; Sequence 5869, Application US/09543681A

; Patent No. 6605709

; GENERAL INFORMATION:

; APPLICANT: GARY BRETON

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS

; FILE REFERENCE: 2709.1002-001

; CURRENT APPLICATION NUMBER: US/09/543,681A

; CURRENT FILING DATE: 2000-04-05

; PRIOR FILING DATE: 1999-04-09

; NUMBER OF SEQ ID NOS: 8344

; SEQ ID NO 5869
; LENGTH: 637
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-5869

Query Match 81.8%; Score 18; DB 4; Length 637;
Best Local Similarity 57.1%; Pred. No. 2.4e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
|
Db 213 VALSAEF 219

RESULT 53

US-09-268-347-46

; Sequence 46, Application US/09268347

; Patent No. 6335182

; GENERAL INFORMATION:

; APPLICANT: Loosmore, Sheena M.

; TITLE OF INVENTION: RECOMBINANT HAEMOPHILUS INFLUENZAE ADHESIN PROTEINS

; FILE REFERENCE: 1038-860

; CURRENT APPLICATION NUMBER: US/09/268,347

; CURRENT FILING DATE: 1999-03-16

; NUMBER OF SEQ ID NOS: 54

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 46

; LENGTH: 659

; TYPE: PRT

; ORGANISM: Haemophilus influenzae

US-09-268-347-46

Query Match 81.8%; Score 18; DB 3; Length 659;
Best Local Similarity 57.1%; Pred. No. 2.5e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
|
Db 573 VASNAEF 579

RESULT 54

US-09-268-347-45

; Sequence 45, Application US/09268347

; Patent No. 6335182

; GENERAL INFORMATION:

; APPLICANT: Loosmore, Sheena M.

; TITLE OF INVENTION: RECOMBINANT HAEMOPHILUS INFLUENZAE ADHESIN PROTEINS

; FILE REFERENCE: 1038-860

; CURRENT APPLICATION NUMBER: US/09/268,347

; CURRENT FILING DATE: 1999-03-16

; NUMBER OF SEQ ID NOS: 54

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 45

; LENGTH: 660

; TYPE: PRT

; ORGANISM: Haemophilus influenzae

US-09-268-347-45

Query Match 81.8%; Score 18; DB 3; Length 660;
Best Local Similarity 57.1%; Pred. No. 2.5e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
|
Db 532 VASNAEF 538

RESULT 55

US-09-252-991A-31488

; Sequence 31488, Application US/09252991A

; Patent No. 6551795

;; GENERAL INFORMATION:
;; APPLICANT: Marc J. Rubenfield et al.
;; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;; FILE REFERENCE: 107196.136
;; CURRENT APPLICATION NUMBER: US/09/252,991A
;; PRIOR FILING DATE: 1999-02-18
;; PRIOR APPLICATION NUMBER: US 60/074,788
;; PRIOR FILING DATE: 1998-02-18
;; PRIOR APPLICATION NUMBER: US 60/094,190
;; PRIOR FILING DATE: 1998-07-27
;; NUMBER OF SEQ ID NOS: 33142
;; SEQ ID NO 31488
;; LENGTH: 669
;; TYPE: PRT
;; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-31488

Query Match 81.8%; Score 18; DB 4; Length 669;
Best Local Similarity 57.1%; Pred. No. 2.6e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VXXAAEF 7
| |||
Db 588 VYAAAEF 594

RESULT 56
US-09-363-708-2
;; Sequence 2, Application US/09363708
;; Patent No. 6399747
;; GENERAL INFORMATION:
;; APPLICANT: Schmandt, et al.
;; TITLE OF INVENTION: NOVEL SHC BINDING PROTEIN
;; NUMBER OF SEQUENCES: 12
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
;; STREET: 233 South Wacker Drive/6300 Sears Tower
;; CITY: Chicago
;; STATE: Illinois
;; COUNTRY: United States of America
;; ZIP: 60606-6402
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/363,708
;; FILING DATE:
;; CLASSIFICATION:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Clough, David W.
;; REGISTRATION NUMBER: 36,107
;; REFERENCE/DOCKET NUMBER: 01017/34451
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (312) 474-6300
;; TELEFAX: (312) 474-0448
;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 672 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: not relevant
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;; DESCRIPTION: /desc = "hPAL peptide"

US-09-363-708-2
Query Match 81.8%; Score 18; DB 3; Length 672;
Best Local Similarity 57.1%; Pred. No. 2.6e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 1 VXXAAEF 7

Db 471 VRTSAEF 477
| |||
RESULT 57
US-09-083-587-2
;; Sequence 2, Application US/09083587
;; Patent No. 6492138
;; GENERAL INFORMATION:
;; APPLICANT: Schmandt, et al.
;; TITLE OF INVENTION: NOVEL SHC BINDING PROTEIN
;; NUMBER OF SEQUENCES: 12
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
;; STREET: 233 South Wacker Drive/6300 Sears Tower
;; CITY: Chicago
;; STATE: Illinois
;; COUNTRY: United States of America
;; ZIP: 60606-6402
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/083,587
;; FILING DATE:
;; CLASSIFICATION:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Clough, David W.
;; REGISTRATION NUMBER: 36,107
;; REFERENCE/DOCKET NUMBER: 01017/34451
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (312) 474-6300
;; TELEFAX: (312) 474-0448
;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 672 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: not relevant
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;; DESCRIPTION: /desc = "hPAL peptide"

US-09-083-587-2
Query Match 81.8%; Score 18; DB 4; Length 672;
Best Local Similarity 57.1%; Pred. No. 2.6e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VXXAAEF 7
| |||
Db 471 VRTSAEF 477

RESULT 58
US-08-772-270A-12
;; Sequence 12, Application US/08772270A
;; Patent No. 6019984
;; GENERAL INFORMATION:
;; APPLICANT: MacInnes, Janet
;; APPLICANT: Ricciatti, Paul
;; APPLICANT: Mallard, Bonnie
;; APPLICANT: Rosendal, Soren
;; TITLE OF INVENTION: NOVEL BACTERIAL PREPARATIONS, METHOD FOR
;; NUMBER OF SEQUENCES: 14
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Bereskin & Parr
;; STREET: 40 King Street West
;; CITY: Toronto
;; STATE: Ontario
;; COUNTRY: Canada
;; ZIP: M5H 3Y2

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION NUMBER: US/08/772,270A
; FILING DATE: December 23, 1996
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Gravelle, Micheline
; REGISTRATION NUMBER: 40,261
; REFERENCE/DOCKET NUMBER: 6580-81
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 364-7311
; TELEFAX: (416) 361-1398
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 711 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Actinobacillus pleuropneumoniae
; US-08-772-270A-12

```

```

Query Match      81.8%; Score 18; DB 3; Length 711;
Best Local Similarity 57.1%; Pred. No. 2.7e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

Qy 1 VXXAAEF 7
| |||
Db 112 VLSAAEF 118

```

```

RESULT 59
US-09-569-037-8
; Sequence 8, Application US/09569037
; Patent No. 6753151
; GENERAL INFORMATION:
; APPLICANT: Jackson, Stephen P
; APPLICANT: Gell, David A
; TITLE OF INVENTION: Interactions of Ku polypeptides and applications
; TITLE OF INVENTION: thereof
; FILE REFERENCE: 620-100
; CURRENT APPLICATION NUMBER: US/09/569,037
; CURRENT FILING DATE: 2000-05-11
; PRIOR APPLICATION NUMBER: PCT/GB00/01245
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: GB 9907687.9
; PRIOR FILING DATE: 1999-04-01
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 8
; LENGTH: 728
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
; US-09-569-037-8

```

```

Query Match      81.8%; Score 18; DB 4; Length 728;
Best Local Similarity 57.1%; Pred. No. 2.8e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

Qy 1 VXXAAEF 7
| |||
Db 718 VANAAEF 724

```

```

RESULT 60
US-07-603-133B-12
; Sequence 12, Application US/07603133B
; Patent No. 5298244

```

```

; GENERAL INFORMATION:
; APPLICANT: Redmond, Mark J.
; APPLICANT: Ijaz, Mohammed K.
; APPLICANT: Parker, Michael D.
; TITLE OF INVENTION: ASSEMBLED VIRAL PARTICLES AND THEIR
; TITLE OF INVENTION: USE IN A VACCINE TO ROTAVIRAL DISEASE
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morrison & Foerster
; STREET: 545 Middlefield Road, Suite 200
; CITY: Menlo Park
; STATE: CA
; COUNTRY: USA
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/603,133B
; FILING DATE: 19901025
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Robins, Roberta L.
; REGISTRATION NUMBER: 33,208
; REFERENCE/DOCKET NUMBER: 9313-0004.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 327-7250
; TELEFAX: (415) 327-2951
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 775 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-07-603-133B-12

```

```

Query Match      81.8%; Score 18; DB 1; Length 775;
Best Local Similarity 57.1%; Pred. No. 3e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

Qy 1 VXXAAEF 7
| |||
Db 199 VSSDAEF 205

```

```

RESULT 61
US-09-248-796A-16379
; Sequence 16379, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 16379
; LENGTH: 1010
; TYPE: PRT
; ORGANISM: Candida albicans
; US-09-248-796A-16379

```

```

Query Match      81.8%; Score 18; DB 4; Length 1010;
Best Local Similarity 57.1%; Pred. No. 4e+03;

```

```
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VXXAEF 7
Db 621 VTTEAEF 627

RESULT 62
PCT-US95-02251-3
; Sequence 3, Application PC/TUS9502251
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR STIMULATING BONE
; TITLE OF INVENTION: CELLS
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: United States of America
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII
; SOFTWARE: PatentIn Release #1.0, Version
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/02251
; FILING DATE: CONCURRENTLY HERewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/316,650
; FILING DATE: 30-SEP-1994
; CLASSIFICATION:
; APPLICATION NUMBER: US 08/199,780
; FILING DATE: 18-FEB-1994
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Parker, David L.
; REGISTRATION NUMBER: 32,165
; REFERENCE/DOCKET NUMBER: UMIC009P--
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (713) 789-2679
; TELEX: 79-0924
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1251 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US95-02251-3

Query Match 81.8%; Score 18; DB 5; Length 1251;
Best Local Similarity 57.1%; Pred. No. 5e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VXXAEF 7
Db 940 VYSSAEF 946

RESULT 63
US-08-199-780-3
; Sequence 3, Application US/08199780
; Patent No. 5763416
; GENERAL INFORMATION:
; APPLICANT: Bonadio, Jeffrey
; APPLICANT: Goldstein, Steven A.
; TITLE OF INVENTION: Gene Transfer Into Bone Cells
; TITLE OF INVENTION: And Tissues
```

```
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: TX
; COUNTRY: USA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30B
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/199,780
; FILING DATE: 18-FEB-1994
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Parker, David
; REGISTRATION NUMBER: 32,165
; REFERENCE/DOCKET NUMBER: UMIC:002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 320-7200
; TELEFAX: (512) 474-7577
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1252 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-199-780-3

Query Match 81.8%; Score 18; DB 1; Length 1252;
Best Local Similarity 57.1%; Pred. No. 5e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VXXAEF 7
Db 941 VYSSAEF 947

RESULT 64
US-08-316-650-3
; Sequence 3, Application US/08316650
; Patent No. 5942496
; GENERAL INFORMATION:
; APPLICANT: Bonadio, Jeffrey
; APPLICANT: Roessler, Blake J.
; APPLICANT: Goldstein, Steven A.
; APPLICANT: Lin, Wushan
; TITLE OF INVENTION: METHODS AND COMPOSITIONS
; TITLE OF INVENTION: FOR STIMULATING BONE CELLS
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/316,650
; FILING DATE: 30-SEP-1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/199,780
; FILING DATE: 30-SEP-1994
; ATTORNEY/AGENT INFORMATION:
```

NAME: Parker, David L.
REGISTRATION NUMBER: 32,165
REFERENCE/DOCKET NUMBER: UMIC:008
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (713) 789-2679
TELEX: 79-0924
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1252 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-316-650-3

Query Match 81.8%; Score 18; DB 2; Length 1252;
Best Local Similarity 57.1%; Pred. No. 5e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
| |||
Db 941 VYSSAEF 947

RESULT 65
US-08-479-722B-4
; Sequence 4, Application US/08479722B
; Patent No. 6074840
; GENERAL INFORMATION:
; APPLICANT: Bonadio, Jeffrey
; APPLICANT: Yin, Wushan
; TITLE OF INVENTION: LATENT TGF(BINDING PROTEIN (LTBP)
; GENES, COMPOSITIONS AND METHODS
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Williams, Morgan & Amerson
; STREET: 7676 Hillmont, Suite 250
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77040
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/479,722B
; FILING DATE: 07-JUN-1995
; CLASSIFICATION:
; PRIOR APPLICATION NUMBER: US PCT/US95/02251
; FILING DATE: 21-FEB-1995
; APPLICATION NUMBER: US 08/316,650
; FILING DATE: 30-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/199,780
; FILING DATE: 18-FEB-1994
; NAME: Fussey, Shelley P.M.
; REGISTRATION NUMBER: 39,458
; REFERENCE/DOCKET NUMBER: 4100.000500/FUS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (713) 934-7000
; TELEFAX: (713) 934-7011
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1253 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-08-479-722B-4

Query Match 81.8%; Score 18; DB 2; Length 1252;
Best Local Similarity 57.1%; Pred. No. 5e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
| |||
Db 941 VYSSAEF 947

RESULT 66
US-09-592-685-4
; Sequence 4, Application US/09592685
; Patent No. 6774105
; GENERAL INFORMATION:
; APPLICANT: Bonadio, Jeffrey
; APPLICANT: Yin, Wushan
; TITLE OF INVENTION: LATENT TGF(BINDING PROTEIN (LTBP)
; GENES, COMPOSITIONS AND METHODS
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Williams, Morgan & Amerson
; STREET: 7676 Hillmont, Suite 250
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77040
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/592,685
; FILING DATE: 12-JUN-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION NUMBER: US/08/479,722B
; FILING DATE: 07-JUN-1995
; APPLICATION NUMBER: US PCT/US95/02251
; FILING DATE: 21-FEB-1995
; APPLICATION NUMBER: US 08/316,650
; FILING DATE: 30-SEP-1994
; APPLICATION NUMBER: US 08/199,780
; FILING DATE: 18-FEB-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Fussey, Shelley P.M.
; REGISTRATION NUMBER: 39,458
; REFERENCE/DOCKET NUMBER: 4100.000500/FUS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (713) 934-7000
; TELEFAX: (713) 934-7011
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1253 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-592-685-4

Query Match 81.8%; Score 18; DB 4; Length 1253;
Best Local Similarity 57.1%; Pred. No. 5e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
| |||
Db 941 VYSSAEF 947

RESULT 67
US-09-077-098A-7
; Sequence 7, Application US/09077098A

NAME: Parker, David L.
REGISTRATION NUMBER: 32,165
REFERENCE/DOCKET NUMBER: UMIC:008
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (713) 789-2679
TELEX: 79-0924
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1252 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-316-650-3

Query Match 81.8%; Score 18; DB 3; Length 1253;
Best Local Similarity 57.1%; Pred. No. 5e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
| |||
Db 941 VYSSAEF 947

RESULT 66
US-09-592-685-4
; Sequence 4, Application US/09592685
; Patent No. 6774105
; GENERAL INFORMATION:
; APPLICANT: Bonadio, Jeffrey
; APPLICANT: Yin, Wushan
; TITLE OF INVENTION: LATENT TGF(BINDING PROTEIN (LTBP)
; GENES, COMPOSITIONS AND METHODS
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Williams, Morgan & Amerson
; STREET: 7676 Hillmont, Suite 250
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77040
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/592,685
; FILING DATE: 12-JUN-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION NUMBER: US/08/479,722B
; FILING DATE: 07-JUN-1995
; APPLICATION NUMBER: US PCT/US95/02251
; FILING DATE: 21-FEB-1995
; APPLICATION NUMBER: US 08/316,650
; FILING DATE: 30-SEP-1994
; APPLICATION NUMBER: US 08/199,780
; FILING DATE: 18-FEB-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Fussey, Shelley P.M.
; REGISTRATION NUMBER: 39,458
; REFERENCE/DOCKET NUMBER: 4100.000500/FUS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (713) 934-7000
; TELEFAX: (713) 934-7011
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1253 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-592-685-4

Query Match 81.8%; Score 18; DB 4; Length 1253;
Best Local Similarity 57.1%; Pred. No. 5e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
| |||
Db 941 VYSSAEF 947

RESULT 67
US-09-077-098A-7
; Sequence 7, Application US/09077098A

```
; Patent No. 6544519
; GENERAL INFORMATION:
; APPLICANT: TOKUNAGA, Ei-ji
; SAKAGUCHI, Masashi
; MATSUO, Kazuo
; HAMADA, Fukusaburo
; TOKIYOSHI, Sachio
; TITLE OF INVENTION: NOVEL POLYPEPTIDE FROM HAEMOPHILUS
; PARAGALLINARUM AND PROCESS FOR PREPARING THE SAME
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 624 Ninth Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/077,098A
; FILING DATE: 19-May-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/JP97/03222
; FILING DATE: 12-SEP-1997
; APPLICATION NUMBER: JP 27,148/1996
; FILING DATE: 19-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: KORNEAU, Anne M.
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: TOKUNAGA=1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2039 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-09-077-098A-7

Query Match      81.8%; Score 18; DB 4; Length 2039;
Best Local Similarity 57.1%; Pred. No. 8.5e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy      1 VXXAAEF 7
Db      422 VAAGAEF 428

RESULT 68
US-09-949-016-10933
; Sequence 10933, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/178,368
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6978
; LENGTH: 4377
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-6978

Query Match      81.8%; Score 18; DB 4; Length 4377;
Best Local Similarity 42.9%; Pred. No. 1.9e+04;
Matches 3; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy      1 VXXAAEF 7
Db      3397 IATTAEF 3403

RESULT 70
US-09-604-608-27
; Sequence 27, Application US/09604608
; Patent No. 6545127
; GENERAL INFORMATION:
; APPLICANT: Tang, Jordan J.N.
; APPLICANT: Lin, Xinli
; APPLICANT: Koelsch, Gerald
; TITLE OF INVENTION: Catalytically Active Recombinant Menapsin and Methods
; FILE REFERENCE: OMPF 179
; CURRENT APPLICATION NUMBER: US/09/604,608
; CURRENT FILING DATE: 2000-06-27
; PRIOR FILING DATE: 2000-06-27
; PRIOR APPLICATION NUMBER: 60/141,363
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: 60/168,060
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: 60/177,836
; PRIOR FILING DATE: 2000-01-25
; PRIOR APPLICATION NUMBER: 60/178,368
; PRIOR FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: 60/210,292
```

```
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10933
; LENGTH: 3913
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-10933

Query Match      81.8%; Score 18; DB 4; Length 3913;
Best Local Similarity 42.9%; Pred. No. 1.7e+04;
Matches 3; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy      1 VXXAAEF 7
Db      2933 IATTAEF 2939

RESULT 69
US-09-949-016-6978
; Sequence 6978, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6978
; LENGTH: 4377
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-6978

Query Match      81.8%; Score 18; DB 4; Length 4377;
Best Local Similarity 42.9%; Pred. No. 1.9e+04;
Matches 3; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy      1 VXXAAEF 7
Db      3397 IATTAEF 3403

RESULT 70
US-09-604-608-27
; Sequence 27, Application US/09604608
; Patent No. 6545127
; GENERAL INFORMATION:
; APPLICANT: Tang, Jordan J.N.
; APPLICANT: Lin, Xinli
; APPLICANT: Koelsch, Gerald
; TITLE OF INVENTION: Catalytically Active Recombinant Menapsin and Methods
; FILE REFERENCE: OMPF 179
; CURRENT APPLICATION NUMBER: US/09/604,608
; CURRENT FILING DATE: 2000-06-27
; PRIOR FILING DATE: 2000-06-27
; PRIOR APPLICATION NUMBER: 60/141,363
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: 60/168,060
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: 60/177,836
; PRIOR FILING DATE: 2000-01-25
; PRIOR APPLICATION NUMBER: 60/178,368
; PRIOR FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: 60/210,292
```


; PRIOR FILING DATE: 2000-06-08
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 27
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Inhibitors
US-09-604-608-27

Query Match 77.3%; Score 17; DB 4; Length 7;
Best Local Similarity 57.1%; Pred. No. 4.1e+05;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VXXAAEF 7
| |||
Db 1 VNLAAEF 7

RESULT 71
US-09-604-608-28
; Sequence 28, Application US/09604608
; Patent No. 6545127
; GENERAL INFORMATION:
; APPLICANT: Tang, Jordan J.N.
; APPLICANT: Lin, Xinli
; APPLICANT: Koelsch, Gerald

; TITLE OF INVENTION: Catalytically Active Recombinant Memapsin and Methods
; FILE REFERENCE: OMRF 179
; CURRENT APPLICATION NUMBER: US/09/604,608
; PRIOR FILING DATE: 2000-06-27
; PRIOR FILING DATE: 1999-06-28
; PRIOR FILING DATE: 1999-11-30
; PRIOR FILING DATE: 2000-01-25
; PRIOR FILING DATE: 2000-01-27
; PRIOR FILING DATE: 2000-06-08
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 28
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Inhibitors
US-09-604-608-28

Query Match 77.3%; Score 17; DB 4; Length 8;
Best Local Similarity 57.1%; Pred. No. 4.1e+05;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VXXAAEF 7
| |||
Db 2 VNLAAEF 8

RESULT 72
US-09-724-566A-87
; Sequence 87, Application US/09724566A
; Patent No. 6627739
; GENERAL INFORMATION:
; APPLICANT: Anderson, John P.
; APPLICANT: Basi, Gurigbal
; APPLICANT: Doane, Minh Tam
; APPLICANT: Frigon, No. 6627739mand
; APPLICANT: John, Varghese
; APPLICANT: Power, Michael

; APPLICANT: Sinha, Sukanto
; APPLICANT: Tatsuno, Gwen
; APPLICANT: Tung, Jay
; APPLICANT: Wang, Shuwen
; APPLICANT: McConlogue, Lisa
; TITLE OF INVENTION: Beta-Secretase Enzyme Compositions and
; FILE REFERENCE: 228-US-NEWC2
; CURRENT FILING DATE: 2000-11-28
; PRIOR FILING DATE: 2000-02-10
; PRIOR FILING DATE: 1999-02-10
; PRIOR FILING DATE: 1999-02-10
; PRIOR FILING DATE: 1999-06-15
; NUMBER OF SEQ ID NOS: 104
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 87
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: APP fragment
US-09-724-566A-87

Query Match 77.3%; Score 17; DB 4; Length 9;
Best Local Similarity 57.1%; Pred. No. 4.1e+05;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VXXAAEF 7
| |||
Db 3 VKMAAEF 9

RESULT 73
US-09-724-566A-88
; Sequence 88, Application US/09724566A
; Patent No. 6627739
; GENERAL INFORMATION:
; APPLICANT: Anderson, John P.
; APPLICANT: Basi, Gurigbal
; APPLICANT: Doane, Minh Tam
; APPLICANT: Frigon, No. 6627739mand
; APPLICANT: John, Varghese
; APPLICANT: Power, Michael
; APPLICANT: Sinha, Sukanto
; APPLICANT: Tatsuno, Gwen
; APPLICANT: Tung, Jay
; APPLICANT: Wang, Shuwen
; APPLICANT: McConlogue, Lisa
; TITLE OF INVENTION: Beta-Secretase Enzyme Compositions and
; FILE REFERENCE: 228-US-NEWC2
; CURRENT FILING DATE: 2000-11-28
; PRIOR FILING DATE: 2000-02-10
; PRIOR FILING DATE: 2000-02-10
; PRIOR FILING DATE: 1999-02-10
; PRIOR FILING DATE: 1999-06-15
; NUMBER OF SEQ ID NOS: 104
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 88
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: APP fragment
US-09-724-566A-88

. Query Match 77.3%; Score 17; DB 4; Length 9;

Best Local Similarity 57.1%; Pred. No. 4.1e+05;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VXXAAEF 7
| |||
Db 3 VNLAAEF 9

RESULT 74

US-09-724-566A-89
; Sequence 89, Application US/09724566A
; Patent No. 6627739
; GENERAL INFORMATION:
; APPLICANT: Anderson, John P.
; APPLICANT: Basi, Gurigbal
; APPLICANT: Doane, Minh Tam
; APPLICANT: Frigon, No. 6627739mand
; APPLICANT: John, Varghese
; APPLICANT: Power, Michael
; APPLICANT: Sinha, Sukanto
; APPLICANT: Tatsuno, Gwen
; APPLICANT: Tung, Jay
; APPLICANT: Wang, Shuwen
; APPLICANT: McConlogue, Lisa
; TITLE OF INVENTION: Beta-Secretase Enzyme Compositions and
; FILE REFERENCE: 228-US-NEWC2
; CURRENT APPLICATION NUMBER: US/09/724,566A
; CURRENT FILING DATE: 2000-11-28
; PRIOR FILING DATE: 2000-02-10
; PRIOR APPLICATION NUMBER: US 09/501,708
; PRIOR FILING DATE: 2000-02-10
; PRIOR APPLICATION NUMBER: 60/119,571
; PRIOR FILING DATE: 1999-02-10
; PRIOR APPLICATION NUMBER: 60/139,172
; PRIOR FILING DATE: 1999-06-15
; NUMBER OF SEQ ID NOS: 104
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 89
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: APP fragment
US-09-724-566A-89

Query Match 77.3%; Score 17; DB 4; Length 9;
Best Local Similarity 57.1%; Pred. No. 4.1e+05;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VXXAAEF 7
| |||
Db 3 VKLAAEF 9

RESULT 75

US-09-724-566A-93
; Sequence 93, Application US/09724566A
; Patent No. 6627739
; GENERAL INFORMATION:
; APPLICANT: Anderson, John P.
; APPLICANT: Basi, Gurigbal
; APPLICANT: Doane, Minh Tam
; APPLICANT: Frigon, No. 6627739mand
; APPLICANT: John, Varghese
; APPLICANT: Power, Michael
; APPLICANT: Sinha, Sukanto
; APPLICANT: Tatsuno, Gwen
; APPLICANT: Tung, Jay
; APPLICANT: Wang, Shuwen
; APPLICANT: McConlogue, Lisa
; TITLE OF INVENTION: Beta-Secretase Enzyme Compositions and
; FILE REFERENCE: 228-US-NEWC2

; CURRENT APPLICATION NUMBER: US/09/724,566A
; CURRENT FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: US 09/501,708
; PRIOR FILING DATE: 2000-02-10
; PRIOR APPLICATION NUMBER: 60/119,571
; PRIOR FILING DATE: 1999-02-10
; PRIOR APPLICATION NUMBER: 60/139,172
; PRIOR FILING DATE: 1999-06-15
; NUMBER OF SEQ ID NOS: 104
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 93
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: APP fragment
US-09-724-566A-93

Query Match 77.3%; Score 17; DB 4; Length 9;
Best Local Similarity 57.1%; Pred. No. 4.1e+05;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VXXAAEF 7
| |||
Db 3 VKFAAEF 9

RESULT 76

US-09-724-566A-94
; Sequence 94, Application US/09724566A
; Patent No. 6627739
; GENERAL INFORMATION:
; APPLICANT: Anderson, John P.
; APPLICANT: Basi, Gurigbal
; APPLICANT: Doane, Minh Tam
; APPLICANT: Frigon, No. 6627739mand
; APPLICANT: John, Varghese
; APPLICANT: Power, Michael
; APPLICANT: Sinha, Sukanto
; APPLICANT: Tatsuno, Gwen
; APPLICANT: Tung, Jay
; APPLICANT: Wang, Shuwen
; APPLICANT: McConlogue, Lisa
; TITLE OF INVENTION: Beta-Secretase Enzyme Compositions and
; FILE REFERENCE: 228-US-NEWC2
; CURRENT APPLICATION NUMBER: US/09/724,566A
; CURRENT FILING DATE: 2000-11-28
; PRIOR FILING DATE: 2000-02-10
; PRIOR APPLICATION NUMBER: US 09/501,708
; PRIOR FILING DATE: 2000-02-10
; PRIOR APPLICATION NUMBER: 60/119,571
; PRIOR FILING DATE: 1999-02-10
; PRIOR APPLICATION NUMBER: 60/139,172
; PRIOR FILING DATE: 1999-06-15
; NUMBER OF SEQ ID NOS: 104
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 94
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: APP fragment
US-09-724-566A-94

Query Match 77.3%; Score 17; DB 4; Length 9;
Best Local Similarity 57.1%; Pred. No. 4.1e+05;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VXXAAEF 7
| |||
Db 3 VNFAAEF 9

```
RESULT 77
US-09-471-669A-87
; Sequence 87, Application US/09471669A
; Patent No. 6830918
; GENERAL INFORMATION:
; APPLICANT: Anderson, John P.
; APPLICANT: Basi, Gurigbal
; APPLICANT: Doane, Minh Tam
; APPLICANT: Frigon, No. 6830918mand
; APPLICANT: John, Varghese
; APPLICANT: Power, Michael
; APPLICANT: Sinha, Sukanto
; APPLICANT: Tatsuno, Gwen
; APPLICANT: Tung, Jay
; APPLICANT: Wang, Shuwen
; APPLICANT: McConlogue, Lisa
; APPLICANT: Elan Pharmaceuticals, Inc.
; TITLE OF INVENTION: BETA-SECRETASE ENZYME COMPOSITIONS AND METHODS
; FILE REFERENCE: 015270-006430US
; CURRENT APPLICATION NUMBER: US/09/471.669A
; CURRENT FILING DATE: 1999-12-24
; PRIOR FILING DATE: 1998-12-31
; PRIOR APPLICATION NUMBER: US 60/114,408
; PRIOR FILING DATE: 1999-02-10
; PRIOR APPLICATION NUMBER: US 60/119,571
; PRIOR FILING DATE: 1999-02-10
; PRIOR APPLICATION NUMBER: US 60/139,172
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 87
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: APP fragment
US-09-471-669A-87

Query Match 77.3%; Score 17; DB 4; Length 9;
Best Local Similarity 57.1%; Pred. No. 4.1e+05;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VXXAAEF 7
| |||
Db 3 VKMAAEF 9

RESULT 78
US-09-471-669A-88
; Sequence 88, Application US/09471669A
; Patent No. 6830918
; GENERAL INFORMATION:
; APPLICANT: Anderson, John P.
; APPLICANT: Basi, Gurigbal
; APPLICANT: Doane, Minh Tam
; APPLICANT: Frigon, No. 6830918mand
; APPLICANT: John, Varghese
; APPLICANT: Power, Michael
; APPLICANT: Sinha, Sukanto
; APPLICANT: Tatsuno, Gwen
; APPLICANT: Tung, Jay
; APPLICANT: Wang, Shuwen
; APPLICANT: McConlogue, Lisa
; APPLICANT: Elan Pharmaceuticals, Inc.
; TITLE OF INVENTION: BETA-SECRETASE ENZYME COMPOSITIONS AND METHODS
; FILE REFERENCE: 015270-006430US
; CURRENT APPLICATION NUMBER: US/09/471.669A
; CURRENT FILING DATE: 1999-12-24
; PRIOR FILING DATE: 1998-12-31
; PRIOR APPLICATION NUMBER: US 60/114,408
; PRIOR FILING DATE: 1999-02-10
; PRIOR APPLICATION NUMBER: US 60/119,571
; PRIOR FILING DATE: 1999-02-10
; PRIOR APPLICATION NUMBER: US 60/139,172
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 88
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: APP fragment
US-09-471-669A-88

Query Match 77.3%; Score 17; DB 4; Length 9;
Best Local Similarity 57.1%; Pred. No. 4.1e+05;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VXXAAEF 7
| |||
Db 3 VKMAAEF 9

RESULT 79
US-09-471-669A-89
; Sequence 89, Application US/09471669A
; Patent No. 6830918
; GENERAL INFORMATION:
; APPLICANT: Anderson, John P.
; APPLICANT: Basi, Gurigbal
; APPLICANT: Doane, Minh Tam
; APPLICANT: Frigon, No. 6830918mand
; APPLICANT: John, Varghese
; APPLICANT: Power, Michael
; APPLICANT: Sinha, Sukanto
; APPLICANT: Tatsuno, Gwen
; APPLICANT: Tung, Jay
; APPLICANT: Wang, Shuwen
; APPLICANT: McConlogue, Lisa
; APPLICANT: Elan Pharmaceuticals, Inc.
; TITLE OF INVENTION: BETA-SECRETASE ENZYME COMPOSITIONS AND METHODS
; FILE REFERENCE: 015270-006430US
; CURRENT APPLICATION NUMBER: US/09/471.669A
; CURRENT FILING DATE: 1999-12-24
; PRIOR FILING DATE: 1998-12-31
; PRIOR APPLICATION NUMBER: US 60/114,408
; PRIOR FILING DATE: 1999-02-10
; PRIOR APPLICATION NUMBER: US 60/119,571
; PRIOR FILING DATE: 1999-02-10
; PRIOR APPLICATION NUMBER: US 60/139,172
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 89
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: APP fragment
US-09-471-669A-89

Query Match 77.3%; Score 17; DB 4; Length 9;
Best Local Similarity 57.1%; Pred. No. 4.1e+05;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VXXAAEF 7
| |||
Db 3 VKLAAEF 9

RESULT 80
US-09-471-669A-93
; Sequence 93, Application US/09471669A
; Patent No. 6830918
; GENERAL INFORMATION:
; APPLICANT: Anderson, John P.
; APPLICANT: Basi, Gurigbal
; APPLICANT: Doane, Minh Tam
; APPLICANT: Frigon, No. 6830918mand
; APPLICANT: John, Varghese
; APPLICANT: Power, Michael
; APPLICANT: Sinha, Sukanto
; APPLICANT: Tatsuno, Gwen
; APPLICANT: Tung, Jay
; APPLICANT: Wang, Shuwen
; APPLICANT: McConlogue, Lisa
; APPLICANT: Elan Pharmaceuticals, Inc.
; TITLE OF INVENTION: BETA-SECRETASE ENZYME COMPOSITIONS AND METHODS
; FILE REFERENCE: 015270-006430US
; CURRENT APPLICATION NUMBER: US/09/471.669A
; CURRENT FILING DATE: 1999-12-24
; PRIOR FILING DATE: 1998-12-31
; PRIOR APPLICATION NUMBER: US 60/114,408
; PRIOR FILING DATE: 1999-02-10
; PRIOR APPLICATION NUMBER: US 60/119,571
; PRIOR FILING DATE: 1999-02-10
; PRIOR APPLICATION NUMBER: US 60/139,172
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 93
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: APP fragment
US-09-471-669A-93

Query Match 77.3%; Score 17; DB 4; Length 9;
Best Local Similarity 57.1%; Pred. No. 4.1e+05;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VXXAAEF 7
| |||
Db 3 VKLAAEF 9
```

```

; APPLICANT: Frigon, No. 6830918mand
; APPLICANT: John, Varghese
; APPLICANT: Power, Michael
; APPLICANT: Sinha, Sukanto
; APPLICANT: Tatsuno, Gwen
; APPLICANT: Tung, Jay
; APPLICANT: Wang, Shuwen
; APPLICANT: McConlogue, Lisa
; APPLICANT: Elan Pharmaceuticals, Inc.
; TITLE OF INVENTION: BETA-SECRETASE ENZYME COMPOSITIONS AND METHODS
; FILE REFERENCE: 015270-006430US
; CURRENT APPLICATION NUMBER: US/09/471.669A
; CURRENT FILING DATE: 1999-12-24
; PRIOR APPLICATION NUMBER: US 60/114,408
; PRIOR FILING DATE: 1998-12-31
; PRIOR APPLICATION NUMBER: US 60/119,571
; PRIOR FILING DATE: 1999-02-10
; PRIOR APPLICATION NUMBER: US 60/139,172
; PRIOR FILING DATE: 1999-06-15
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 93
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: APP fragment
US-09-471-669A-93

Query Match          77.3%; Score 17; DB 4; Length 9;
Best Local Similarity 57.1%; Pred. No. 4.1e+05;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 VXXAAEF 7
        |   |||
Db       3 VKFAAEF 9

RESULT 81
US-09-471-669A-94
; Sequence 94, Application US/09471669A
; Patent No. 6830918
; GENERAL INFORMATION:
; APPLICANT: Anderson, John P.
; APPLICANT: Basi, Guribbal
; APPLICANT: Doane, Minh Tam
; APPLICANT: Frigon, No. 6830918mand
; APPLICANT: John, Varghese
; APPLICANT: Power, Michael
; APPLICANT: Sinha, Sukanto
; APPLICANT: Tatsuno, Gwen
; APPLICANT: Tung, Jay
; APPLICANT: Wang, Shuwen
; APPLICANT: McConlogue, Lisa
; APPLICANT: Elan Pharmaceuticals, Inc.
; TITLE OF INVENTION: BETA-SECRETASE ENZYME COMPOSITIONS AND METHODS
; FILE REFERENCE: 015270-006430US
; CURRENT APPLICATION NUMBER: US/09/471.669A
; CURRENT FILING DATE: 1999-12-24
; PRIOR APPLICATION NUMBER: US 60/114,408
; PRIOR FILING DATE: 1998-12-31
; PRIOR APPLICATION NUMBER: US 60/119,571
; PRIOR FILING DATE: 1999-02-10
; PRIOR APPLICATION NUMBER: US 60/139,172
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 94
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: APP fragment

```

```

US-09-471-669A-94

Query Match          77.3%; Score 17; DB 4; Length 9;
Best Local Similarity 57.1%; Pred. No. 4.1e+05;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 VXXAAEF 7
        |   |||
Db       3 VNFAAEF 9

RESULT 82
US-09-634-955B-24
; Sequence 24, Application US/09634955B
; Patent No. 6511834
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel
; APPLICANT: Cook, William James
; TITLE OF INVENTION: 32142, 21481, 25964, 21686, NOVEL HUMAN DEHYDROGENASE
; FILE REFERENCE: MNI-134
; CURRENT APPLICATION NUMBER: US/09/634,955B
; CURRENT FILING DATE: 2000-08-08
; PRIOR APPLICATION NUMBER: 60/192,002
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.
; SEQ ID NO 24
; LENGTH: 23
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: 3-beta hydroxysteroid dehydrogenase domain
US-09-634-955B-24

Query Match          77.3%; Score 17; DB 4; Length 23;
Best Local Similarity 57.1%; Pred. No. 1.3e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 VXXAAEF 7
        |   |||
Db       4 VTGAAEF 10

RESULT 83
US-08-415-751-12
; Sequence 12, Application US/08415751
; Patent No. 5643772
; GENERAL INFORMATION:
; APPLICANT: PETERSEN, CAROLYN
; APPLICANT: LEECH, JAMES
; APPLICANT: NELSON, RICHARD, C.
; APPLICANT: GUT, JIRI
; TITLE OF INVENTION: POLYPEPTIDES BINDING ANTI-
; TITLE OF INVENTION: CRYPTOSPORIDIUM ANTIBODIES, DNA
; TITLE OF INVENTION: AND RNA ENCODING THEM, HYBRID
; TITLE OF INVENTION: VECTOR AND TRANSFORMED HOST AND
; TITLE OF INVENTION: METHODS FOR IMMUNOTHERAPY AND
; TITLE OF INVENTION: DIAGNOSIS AND KIT
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PHILLIPS, MOORE, LEMPPIO & FINLEY
; STREET: 385 Sherman Avenue, Suite 6
; CITY: Palo Alto
; STATE: California
; COUNTRY: United States of America
; ZIP: 94306-1840
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette - 3.5 inch, 1.44 Kb storage
; COMPUTER: PC
; OPERATING SYSTEM: DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:

```

APPLICATION NUMBER: US/08/415,751
FILING DATE: 03-APR-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/071,880
FILING DATE: June 1, 1993
APPLICATION NUMBER: 07/891,301
FILING DATE: May 29, 1992
ATTORNEY/AGENT INFORMATION:
NAME: Hana Dolezalova
REGISTRATION NUMBER: 30,518
REFERENCE/DOCKET NUMBER: 480.19-2 (HHD)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-1677
TELEFAX: (415) 324-1678
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 47 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: *Cryptosporidium parvum*
FEATURE:
NAME/KEY: Positions coded by nonsense codons are
NAME/KEY: identified as Xaa.
US-08-415-751-12

Query Match 77.3% Score 17; DB 1; Length 47;
Best Local Similarity 57.1% Pred. No. 2.9e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VXXAAEF 7
| | | | |
Db 7 VELAAEF 13

RESULT 84
US-09-248-796A-24127
Sequence 24127, Application US/09248796A
Patent No. 6747137
GENERAL INFORMATION:
APPLICANT: Keith Weinstock et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
FILE REFERENCE: 107196.132
CURRENT APPLICATION NUMBER: US/09/248,796A
CURRENT FILING DATE: 1999-02-12
PRIOR APPLICATION NUMBER: US 60/074,725
PRIOR FILING DATE: 1998-02-13
PRIOR APPLICATION NUMBER: US 60/096,409
PRIOR FILING DATE: 1998-08-13
NUMBER OF SEQ ID NOS: 28208
SEQ ID NO 24127
LENGTH: 66
TYPE: PRT
ORGANISM: *Candida albicans*
US-09-248-796A-24127

Query Match 77.3% Score 17; DB 4; Length 66;
Best Local Similarity 57.1% Pred. No. 4.1e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VXXAAEF 7
| | | | |
Db 54 VKIAAEF 60

RESULT 85
US-09-543-681A-7973
Sequence 7973, Application US/09543681A
Patent No. 6605709

GENERAL INFORMATION:
APPLICANT: GARY BRETON
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 2709.1002-001
CURRENT APPLICATION NUMBER: US/09/543,681A
CURRENT FILING DATE: 2000-04-05
PRIOR APPLICATION NUMBER: US 60/128,706
PRIOR FILING DATE: 1999-04-09
NUMBER OF SEQ ID NOS: 8344
SEQ ID NO 7973
LENGTH: 68
TYPE: PRT
ORGANISM: *Proteus mirabilis*
US-09-543-681A-7973

Query Match 77.3% Score 17; DB 4; Length 68;
Best Local Similarity 57.1% Pred. No. 4.3e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VXXAAEF 7
| | | | |
Db 28 VNAQAEF 34

RESULT 86
5177197-52
Patent No. 5177197
APPLICANT: KANZAKI, TETSUTO; OLOFSSON, ANDERS; MOREN, ANITA;
WERNSTEDT, CHRISTER; HELLMAN, ULF; MIYAZONO, KOHEI; CLAESSON-WELSH,
LENA; HELDIN, CARL-HENRIK
TITLE OF INVENTION: ISOLATED NUCLEOTIDE SEQUENCE EXPRESSING
HUMAN TRANSFORMING GROWTH FACTOR-BETA1-BINDING PROTEIN
NUMBER OF SEQUENCES: 53
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/487,343
FILING DATE: 27-FEB-1990
SEQ ID NO: 52
LENGTH: 68
5177197-52

Query Match 77.3% Score 17; DB 6; Length 68;
Best Local Similarity 57.1% Pred. No. 4.3e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VXXAAEF 7
| | | | |
Db 48 VLGTAEF 54

RESULT 87
5177197-52
Patent No. 5177197
APPLICANT: KANZAKI, TETSUTO; OLOFSSON, ANDERS; MOREN, ANITA;
WERNSTEDT, CHRISTER; HELLMAN, ULF; MIYAZONO, KOHEI; CLAESSON-WELSH,
LENA; HELDIN, CARL-HENRIK
TITLE OF INVENTION: ISOLATED NUCLEOTIDE SEQUENCE EXPRESSING
HUMAN TRANSFORMING GROWTH FACTOR-BETA1-BINDING PROTEIN
NUMBER OF SEQUENCES: 53
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/487,343
FILING DATE: 27-FEB-1990
SEQ ID NO: 52
LENGTH: 68
5177197-52

Query Match 77.3% Score 17; DB 6; Length 68;
Best Local Similarity 57.1% Pred. No. 4.3e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VXXAAEF 7
| | | | |
Db 48 VLGTAEF 54

```
RESULT 88
US-09-248-796A-16636
; Sequence 16636, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 16636
; LENGTH: 74
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-16636

Query Match          77.3%; Score 17; DB 4; Length 74;
Best Local Similarity 57.1%; Pred. No. 4.7e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 VXXAAEF 7
Db      35 VMKAAEF 41

RESULT 89
US-08-311-731A-79
; Sequence 79, Application US/08311731A
; Patent No. 6583266
; GENERAL INFORMATION:
; APPLICANT: SMITH, DOUGLAS
; APPLICANT: MAO, JEN-I
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; TITLE OF INVENTION: RELATING TO MYCOBACTERIUM TUBERCULOSIS AND LAPRAE FOR
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 411
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: WOLF, GREENFIELD & SACKS, P.C.
; STREET: 600 ATLANTIC AVENUE
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: USA
; ZIP: 02210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/311,731A
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: GATES, EDWARD R.
; REGISTRATION NUMBER: 31,616
; REFERENCE/DOCKET NUMBER: C0044/7125
; TELEPHONE: 617/720-3500
; TELEFAX: 617/720-2441
; INFORMATION FOR SEQ ID NO: 79:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 95 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein

US-09-248-796A-16636

Query Match          77.3%; Score 17; DB 4; Length 74;
Best Local Similarity 57.1%; Pred. No. 4.7e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 VXXAAEF 7
Db      35 VMKAAEF 41

RESULT 89
US-08-311-731A-79
; Sequence 79, Application US/08311731A
; Patent No. 6583266
; GENERAL INFORMATION:
; APPLICANT: SMITH, DOUGLAS
; APPLICANT: MAO, JEN-I
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; TITLE OF INVENTION: RELATING TO MYCOBACTERIUM TUBERCULOSIS AND LAPRAE FOR
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 411
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: WOLF, GREENFIELD & SACKS, P.C.
; STREET: 600 ATLANTIC AVENUE
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: USA
; ZIP: 02210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/311,731A
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: GATES, EDWARD R.
; REGISTRATION NUMBER: 31,616
; REFERENCE/DOCKET NUMBER: C0044/7125
; TELEPHONE: 617/720-3500
; TELEFAX: 617/720-2441
; INFORMATION FOR SEQ ID NO: 79:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 95 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
```

```
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: MYCOBACTERIUM LEPRAE
US-08-311-731A-79

Query Match          77.3%; Score 17; DB 4; Length 95;
Best Local Similarity 57.1%; Pred. No. 6.1e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 VXXAAEF 7
Db      36 VRDAAEF 42

RESULT 90
US-08-479-078-27
; Sequence 27, Application US/08479078
; Patent No. 5814466
; GENERAL INFORMATION:
; APPLICANT: Pawson, Anthony
; TITLE OF INVENTION: Method for Assaying for a Substance that
; TITLE OF INVENTION: Affects an SH2-Phosphorylated Ligand Regulatory System
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bereskin & Parr
; STREET: 40 King Street, West
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5H 3Y2
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/479,078
; FILING DATE: June 6, 1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Linda M. Kurdydyk
; REGISTRATION NUMBER: 34,971
; REFERENCE/DOCKET NUMBER: 3153-154
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 364-7311
; TELEFAX: (416) 361-1398
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 96 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
US-08-479-078-27

Query Match          77.3%; Score 17; DB 2; Length 96;
Best Local Similarity 57.1%; Pred. No. 6.2e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 VXXAAEF 7
Db      29 VKDAAEF 35

RESULT 91
US-08-311-731A-185
; Sequence 185, Application US/08311731A
; Patent No. 6583266
; GENERAL INFORMATION:
; APPLICANT: SMITH, DOUGLAS
; APPLICANT: MAO, JEN-I
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; TITLE OF INVENTION: RELATING TO MYCOBACTERIUM TUBERCULOSIS AND LAPRAE FOR
```

;; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
;; NUMBER OF SEQUENCES: 411
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: WOLF GREENFIELD & SACKS, P.C.
;; STREET: 600 ATLANTIC AVENUE
;; CITY: BOSTON
;; STATE: MASSACHUSETTS
;; COUNTRY: USA
;; ZIP: 02210
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patent In Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/311,731A
;; FILING DATE:
;; CLASSIFICATION: 530
;; ATTORNEY/AGENT INFORMATION:
;; NAME: GATES, EDWARD R.
;; REGISTRATION NUMBER: 31,616
;; REFERENCE/DOCKET NUMBER: C0044/7125
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 617/720-3500
;; TELEFAX: 617/720-2441
;; INFORMATION FOR SEQ ID NO: 185:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 112 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; HYPOTHETICAL: YES
;; ORIGINAL SOURCE:
;; ORGANISM: Mycobacterium leprae
;; US-08-311-731A-185

Query Match 77.3%; Score 17; DB 4; Length 112;
Best Local Similarity 57.1%; Pred. No. 7.4e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAEF 7
| |||
Db 53 VRDAEF 59

RESULT 92
US-09-107-532A-5330
; Sequence 5330, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571

;; FILING DATE: July 2, 1997
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Ariniello, Pamela Deneke
;; REGISTRATION NUMBER: 40,489
;; REFERENCE/DOCKET NUMBER: GTC-012
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (781)893-5007
;; TELEFAX: (781)893-8277
;; INFORMATION FOR SEQ ID NO: 5330:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 122 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; HYPOTHETICAL: YES
;; ORIGINAL SOURCE:
;; ORGANISM: Enterococcus faecium
;; FEATURE:
;; NAME/KEY: misc feature
;; LOCATION: (B) LOCATION 1...122
;; SEQUENCE DESCRIPTION: SEQ ID NO: 5330:
US-09-107-532A-5330

Query Match 77.3%; Score 17; DB 4; Length 122;
Best Local Similarity 42.9%; Pred. No. 8.1e+02;
Matches 3; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAEF 7
| |||
Db 18 ITDTAEF 24

RESULT 93
US-09-183-861-49
; Sequence 49, Application US/09183861
; Patent No. 6365165
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Webb, John R.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Skeiky, Yasir A.W.
; TITLE OF INVENTION: LEISHMANIA ANTIGENS FOR USEIN THE THERAPY AND
; NUMBER OF SEQUENCES: 87
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/183,861
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/022,765
; FILING DATE: 12-FEB-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.420C3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 49:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 136 amino acids

```
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Leishmania chagasi
US-09-183-861-49

Query Match 77.3%; Score 17; DB 3; Length 136;
Best Local Similarity 57.1%; Pred. No. 9.1e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
Db 12 VREADEF 18

RESULT 94
US-09-022-765-49
; Sequence 49, Application US/09022765
; Patent No. 6375955
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Webb, John R.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Skeiky, Yasir A.W.
; TITLE OF INVENTION: LEISHMANIA ANTIGENS FOR USEIN THE THERAPY AND
; NUMBER OF SEQUENCES: 87
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/022,765
; FILING DATE: 12-FEB-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.420C3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 49:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 136 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Leishmania chagasi
US-09-022-765-49

Query Match 77.3%; Score 17; DB 3; Length 136;
Best Local Similarity 57.1%; Pred. No. 9.1e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
Db 12 VREADEF 18

RESULT 95
```

```
US-09-551-974A-49
; Sequence 49, Application US/09551974A
; Patent No. 6500437
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Webb, John R.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Skeiky, Yasir A.W.
; TITLE OF INVENTION: LEISHMANIA ANTIGENS FOR USE IN THE
; FILE REFERENCE: 210121.420C5
; CURRENT APPLICATION NUMBER: US/09/551,974A
; CURRENT FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 101
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 49
; LENGTH: 136
; TYPE: PRT
; ORGANISM: Leishmania chagasi
US-09-551-974A-49

Query Match 77.3%; Score 17; DB 4; Length 136;
Best Local Similarity 57.1%; Pred. No. 9.1e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
Db 12 VREADEF 18

RESULT 96
US-09-565-501A-49
; Sequence 49, Application US/09565501A
; Patent No. 6607731
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Webb, John R.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Bhatia, Ajay
; APPLICANT: Coler, Rhea
; APPLICANT: Peter, Probst
; TITLE OF INVENTION: LEISHMANIA ANTIGENS FOR USE IN THE
; FILE REFERENCE: 210121.420C6
; CURRENT APPLICATION NUMBER: US/09/565,501A
; CURRENT FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 49
; LENGTH: 136
; TYPE: PRT
; ORGANISM: Leishmania chagasi
US-09-565-501A-49

Query Match 77.3%; Score 17; DB 4; Length 136;
Best Local Similarity 57.1%; Pred. No. 9.1e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
Db 12 VREADEF 18

RESULT 97
US-09-639-206A-49
; Sequence 49, Application US/09639206A
; Patent No. 6613337
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Campos-Neto, Antonio
```


; APPLICANT: Webb, John R.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Bhatia, Ajay
; APPLICANT: Coler, Rhea
; APPLICANT: Probst, Peter
; TITLE OF INVENTION: LEISHMANIA ANTIGENS FOR USE IN THE
; TITLE OF INVENTION: THERAPY AND DIAGNOSIS OF LEISHMANIASIS
; FILE REFERENCE: 210121.420C7
; CURRENT APPLICATION NUMBER: US/09/639,206A
; CURRENT FILING DATE: 2000-08-14
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 49
; LENGTH: 136
; TYPE: PRT
; ORGANISM: Leishmania chagasi
US-09-639-206A-49

Query Match 77.3%; Score 17; DB 4; Length 136;
Best Local Similarity 57.1%; Pred. No. 9.1e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
| |||
Db 12 VREAAEF 18

RESULT 98

US-09-874-923-49

; Sequence 49, Application US/09874923

; Patent No. 6638517

; GENERAL INFORMATION:

; APPLICANT: Reed, Steven G.

; APPLICANT: Campos-Neco, Antonio

; APPLICANT: Webb, John R.

; APPLICANT: Dillon, Davin C.

; APPLICANT: Skeiky, Yasir A.W.

; APPLICANT: Bhatia, Ajay

; APPLICANT: Coler, Rhea

; APPLICANT: Probst, Peter

; APPLICANT: Brannon, Mark

; TITLE OF INVENTION: LEISHMANIA ANTIGENS FOR USE IN THE

; TITLE OF INVENTION: THERAPY AND DIAGNOSIS OF LEISHMANIASIS

; FILE REFERENCE: 210121.420C8

; CURRENT APPLICATION NUMBER: US/09/874,923

; CURRENT FILING DATE: 2001-06-04

; NUMBER OF SEQ ID NOS: 122

; SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO 49

; LENGTH: 136

; TYPE: PRT

; ORGANISM: Leishmania chagasi

US-09-874-923-49

Query Match 77.3%; Score 17; DB 4; Length 136;
Best Local Similarity 57.1%; Pred. No. 9.1e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
| |||
Db 12 VREAAEF 18

RESULT 99

US-09-134-000C-4544

; Sequence 4544, Application US/09134000C

; Patent No. 6617156

; GENERAL INFORMATION:

; APPLICANT: Lynn Doucette-Stamm et al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO

; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 032796-032

; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4544
; LENGTH: 143
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-134-000C-4544

Query Match 77.3%; Score 17; DB 4; Length 143;
Best Local Similarity 57.1%; Pred. No. 9.6e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
| |||
Db 89 VSDIAEF 95

RESULT 100

US-09-248-796A-14343

; Sequence 14343, Application US/09248796A

; Patent No. 6747137

; GENERAL INFORMATION:

; APPLICANT: Keith Weinstock et al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS

; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.132

; CURRENT APPLICATION NUMBER: US/09/248,796A

; CURRENT FILING DATE: 1999-02-12

; PRIOR APPLICATION NUMBER: US 60/074,725

; PRIOR FILING DATE: 1998-02-13

; PRIOR APPLICATION NUMBER: US 60/096,409

; PRIOR FILING DATE: 1998-08-13

; NUMBER OF SEQ ID NOS: 28208

; SEQ ID NO 14343

; LENGTH: 143

; TYPE: PRT

; ORGANISM: Candida albicans

US-09-248-796A-14343

Query Match 77.3%; Score 17; DB 4; Length 143;
Best Local Similarity 57.1%; Pred. No. 9.6e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
| |||
Db 125 VKKTAEF 131

Search completed: June 13, 2005, 14:00:46

Job time : 34 secs

THIS PAGE LEFT BLANK

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2005, 14:00:08 ; Search time 109 Seconds
(without alignments)
24.618 Million cell updates/sec

Title: 09730329-59ED

Perfect score: 22

Sequence: 1 VXXAAEF 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1710399 seqs, 383334425 residues

Total number of hits satisfying chosen parameters: 1710399

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 300 summaries

Database :

Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/PCT NEW PUB.pap.*
- 2: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pap.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW PUB.pap.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pap.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW PUB.pap.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pap.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW PUB.pap.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pap.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pap.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pap.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pap.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW PUB.pap.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pap.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pap.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pap.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pap.*
- 17: /cgn2_6/ptodata/1/pubpaa/US10E_PUBCOMB.pap.*
- 18: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pap.*
- 19: /cgn2_6/ptodata/1/pubpaa/US11A_PUBCOMB.pap.*
- 20: /cgn2_6/ptodata/1/pubpaa/US11_NEW PUB.pap.*
- 21: /cgn2_6/ptodata/1/pubpaa/US60_NEW PUB.pap.*
- 22: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	ID	Description
1	19	86.4	78	US-10-437-963-200225
2	19	86.4	130	US-10-437-963-161742
3	19	86.4	146	US-10-437-963-126883
4	19	86.4	241	US-10-425-115-361705
5	19	86.4	267	US-09-911-781-24
6	19	86.4	267	US-09-976-800-106
7	19	86.4	267	US-10-138-838-106
8	19	86.4	267	US-10-139-031-106
9	19	86.4	267	US-10-138-905-106
10	19	86.4	267	US-10-138-916-106
11	19	86.4	267	US-10-139-296-106
12	19	86.4	267	US-10-437-963-165533
13	19	86.4	267	US-10-437-963-121186
14	19	86.4	267	US-10-424-599-230315
15	19	86.4	267	US-10-424-599-149058
16	19	86.4	267	US-10-425-115-199685
17	19	86.4	267	US-10-389-566-1581
18	19	86.4	267	US-10-424-599-234753
19	19	86.4	267	US-10-425-115-199164
20	19	86.4	267	US-10-425-115-199164
21	19	86.4	267	US-10-425-115-199164
22	19	86.4	267	US-10-425-115-199164
23	19	86.4	267	US-10-425-115-199164
24	19	86.4	267	US-10-425-115-199164
25	19	86.4	267	US-10-425-115-199164
26	19	86.4	267	US-10-425-115-199164
27	19	86.4	267	US-10-425-115-199164
28	19	86.4	267	US-10-425-115-199164
29	19	86.4	267	US-10-425-115-199164
30	19	86.4	267	US-10-425-115-199164
31	19	86.4	267	US-10-425-115-199164
32	19	86.4	267	US-10-425-115-199164
33	19	86.4	267	US-10-425-115-199164
34	19	86.4	267	US-10-425-115-199164
35	19	86.4	267	US-10-425-115-199164
36	19	86.4	267	US-10-425-115-199164
37	19	86.4	267	US-10-425-115-199164
38	19	86.4	267	US-10-425-115-199164
39	19	86.4	267	US-10-425-115-199164
40	19	86.4	267	US-10-425-115-199164
41	19	86.4	267	US-10-425-115-199164
42	19	86.4	267	US-10-425-115-199164
43	19	86.4	267	US-10-425-115-199164
44	19	86.4	267	US-10-425-115-199164
45	19	86.4	267	US-10-425-115-199164
46	19	86.4	267	US-10-425-115-199164
47	19	86.4	267	US-10-425-115-199164
48	19	86.4	267	US-10-425-115-199164
49	19	86.4	267	US-10-425-115-199164
50	19	86.4	267	US-10-425-115-199164
51	19	86.4	267	US-10-425-115-199164
52	19	86.4	267	US-10-425-115-199164
53	19	86.4	267	US-10-425-115-199164
54	19	86.4	267	US-10-425-115-199164
55	19	86.4	267	US-10-425-115-199164
56	19	86.4	267	US-10-425-115-199164
57	19	86.4	267	US-10-425-115-199164
58	19	86.4	267	US-10-425-115-199164
59	19	86.4	267	US-10-425-115-199164
60	19	86.4	267	US-10-425-115-199164
61	19	86.4	267	US-10-425-115-199164
62	19	86.4	267	US-10-425-115-199164
63	19	86.4	267	US-10-425-115-199164
64	19	86.4	267	US-10-425-115-199164
65	19	86.4	267	US-10-425-115-199164
66	19	86.4	267	US-10-425-115-199164
67	19	86.4	267	US-10-425-115-199164
68	19	86.4	267	US-10-425-115-199164
69	19	86.4	267	US-10-425-115-199164
70	19	86.4	267	US-10-425-115-199164
71	19	86.4	267	US-10-425-115-199164
72	19	86.4	267	US-10-425-115-199164
73	19	86.4	267	US-10-425-115-199164
74	19	86.4	267	US-10-425-115-199164
75	19	86.4	267	US-10-425-115-199164
76	19	86.4	267	US-10-425-115-199164
77	19	86.4	267	US-10-425-115-199164
78	19	86.4	267	US-10-425-115-199164
79	19	86.4	267	US-10-425-115-199164
80	19	86.4	267	US-10-425-115-199164
81	19	86.4	267	US-10-425-115-199164
82	19	86.4	267	US-10-425-115-199164
83	19	86.4	267	US-10-425-115-199164
84	19	86.4	267	US-10-425-115-199164

267	14	US-10-139-218-106	Sequence 106, App
267	14	US-10-400-902-24	Sequence 24, Appl
267	14	US-10-405-660-106	Sequence 106, App
267	15	US-10-138-898-106	Sequence 106, App
298	9	US-09-815-242-10242	Sequence 10242, A
298	15	US-10-282-122A-56631	Sequence 56631, A
330	16	US-10-425-115-302302	Sequence 302302, A
338	15	US-10-282-122A-72638	Sequence 72638, A
355	17	US-10-888-656-2	Sequence 2, Appli
355	17	US-10-425-111-57008	Sequence 57008, A
586	16	US-10-437-963-128132	Sequence 128132, A
747	16	US-10-425-115-336689	Sequence 336689, A
751	15	US-10-425-114-51691	Sequence 51691, A
751	15	US-10-425-114-62401	Sequence 62401, A
1266	15	US-10-282-122A-75669	Sequence 75669, A
1584	17	US-10-489-740-173	Sequence 173, App
2214	15	US-10-369-493-6737	Sequence 6737, App
12	14	US-10-032-818-48	Sequence 48, Appl
12	14	US-10-032-818-49	Sequence 49, Appl
12	14	US-10-032-818-54	Sequence 54, Appl
49	15	US-10-424-599-212707	Sequence 212707, A
56	11	US-09-984-429-158	Sequence 158, App
56	14	US-10-150-111-158	Sequence 158, App
60	16	US-10-425-115-237245	Sequence 237245, A
61	16	US-10-425-115-243983	Sequence 243983, A
66	16	US-10-425-115-275053	Sequence 275053, A
69	15	US-10-334-561A-4	Sequence 4, Appli
86	15	US-10-424-599-228254	Sequence 228254, A
91	15	US-10-425-114-49255	Sequence 49255, A
93	15	US-10-369-493-10066	Sequence 10066, A
95	16	US-10-767-701-53633	Sequence 53633, A
100	16	US-10-425-115-255955	Sequence 255955, A
102	16	US-10-425-115-190652	Sequence 190652, A
116	16	US-10-425-115-368590	Sequence 368590, A
121	15	US-10-424-599-146333	Sequence 146333, A
122	16	US-10-437-963-184577	Sequence 184577, A
122	16	US-10-425-115-271612	Sequence 271612, A
122	16	US-10-425-115-274994	Sequence 274994, A
124	15	US-10-424-599-156801	Sequence 156801, A
124	16	US-10-767-701-33440	Sequence 33440, A
124	16	US-10-425-115-279769	Sequence 279769, A
125	16	US-10-425-115-238741	Sequence 238741, A
129	15	US-10-424-599-228146	Sequence 228146, A
137	15	US-10-424-599-274885	Sequence 274885, A
137	16	US-10-370-715B-390	Sequence 390, App
141	15	US-10-425-111-61704	Sequence 61704, A
142	16	US-10-425-115-274236	Sequence 274236, A
144	15	US-10-424-599-279916	Sequence 279916, A
149	16	US-10-437-963-169685	Sequence 169685, A
155	16	US-10-437-963-198547	Sequence 198547, A
159	15	US-10-369-493-16932	Sequence 16932, A
161	15	US-10-424-599-234753	Sequence 234753, A
165	15	US-10-424-599-245373	Sequence 245373, A
167	15	US-10-425-114-64341	Sequence 64341, A
171	14	US-10-156-761-7611	Sequence 7611, App
179	15	US-10-264-213-199	Sequence 199, App
183	16	US-10-425-115-199164	Sequence 199164, A
187	16	US-10-767-701-40084	Sequence 40084, A
188	15	US-10-424-599-162790	Sequence 162790, A
188	16	US-10-425-115-241928	Sequence 241928, A
194	14	US-10-156-761-9937	Sequence 9937, App
199	16	US-10-425-115-271613	Sequence 271613, A
201	16	US-10-437-963-165533	Sequence 165533, A
209	16	US-10-437-963-121186	Sequence 121186, A
212	15	US-10-424-599-230315	Sequence 230315, A
213	15	US-10-424-599-149058	Sequence 149058, A
213	15	US-10-424-599-149059	Sequence 149059, A
220	16	US-10-425-115-199685	Sequence 199685, A
228	15	US-10-389-566-1581	Sequence 1581, App
228	15	US-10-389-566-1597	Sequence 1597, App
234	15	US-10-424-599-247828	Sequence 247828, A
240	17	US-10-893-671-98	Sequence 98, Appl
241	15	US-10-424-599-262409	Sequence 262409, A

85	18	81.8	252	16	US-10-437-963-115265	Sequence 115265,	158	18	81.8	513	14	US-10-122-026-19	Sequence 19, Appl
86	18	81.8	253	17	US-10-762-107-46	Sequence 46, Appl	159	18	81.8	513	14	US-10-164-279-33	Sequence 33, Appl
87	18	81.8	256	15	US-10-424-599-274705	Sequence 274705,	160	18	81.8	513	15	US-10-366-345-38	Sequence 38, Appl
88	18	81.8	258	9	US-09-738-626-5807	Sequence 5807, Ap	161	18	81.8	513	15	US-10-291-265-816	Sequence 816, App
89	18	81.8	258	16	US-10-781-014-600	Sequence 600, App	162	18	81.8	513	15	US-10-375-150-8	Sequence 8, Appl
90	18	81.8	259	16	US-10-437-963-114362	Sequence 114362,	163	18	81.8	514	16	US-10-425-115-234473	Sequence 234473,
91	18	81.8	268	15	US-10-282-122A-65087	Sequence 65087, A	164	18	81.8	515	15	US-10-423-114-52765	Sequence 52765, A
92	18	81.8	268	15	US-10-282-122A-66020	Sequence 66020, A	165	18	81.8	517	15	US-10-424-599-231340	Sequence 231340,
93	18	81.8	268	15	US-10-275-026A-170	Sequence 170, App	166	18	81.8	516	9	US-09-801-368-70	Sequence 70, Appl
94	18	81.8	271	15	US-10-236-031B-16	Sequence 16, Appl	167	18	81.8	526	15	US-10-369-493-21912	Sequence 21912, A
95	18	81.8	288	15	US-10-282-122A-72181	Sequence 72181, A	168	18	81.8	532	9	US-09-828-302-12	Sequence 12, Appl
96	18	81.8	293	16	US-10-767-701-34443	Sequence 34443, A	169	18	81.8	548	14	US-10-156-761-13442	Sequence 13442, A
97	18	81.8	296	15	US-10-425-114-38227	Sequence 38227, A	170	18	81.8	560	15	US-10-094-749-3278	Sequence 3278, Ap
98	18	81.8	301	15	US-10-424-599-265888	Sequence 265888,	171	18	81.8	566	15	US-10-425-114-66755	Sequence 66755, A
99	18	81.8	312	15	US-10-369-493-9623	Sequence 9623, Ap	172	18	81.8	567	15	US-10-369-493-4028	Sequence 4028, Ap
100	18	81.8	315	15	US-10-282-122A-62445	Sequence 62445, A	173	18	81.8	576	16	US-10-437-963-171121	Sequence 171121,
101	18	81.8	316	15	US-10-369-493-13085	Sequence 13085, A	174	18	81.8	592	15	US-10-389-566-1883	Sequence 1883, Ap
102	18	81.8	323	15	US-10-282-122A-69408	Sequence 69408, A	175	18	81.8	605	15	US-10-389-566-997	Sequence 997, App
103	18	81.8	324	15	US-10-282-122A-67010	Sequence 67010, A	176	18	81.8	605	15	US-10-389-566-1385	Sequence 1385, Ap
104	18	81.8	325	17	US-10-732-923-1378	Sequence 1378, Ap	177	18	81.8	607	9	US-09-344-882-18	Sequence 18, Appl
105	18	81.8	333	15	US-10-282-122A-61742	Sequence 61742, A	178	18	81.8	607	14	US-10-293-865-18	Sequence 18, Appl
106	18	81.8	333	15	US-10-282-122A-62445	Sequence 62445, A	179	18	81.8	607	15	US-10-389-566-553	Sequence 553, App
107	18	81.8	333	15	US-10-282-122A-63915	Sequence 63915, A	180	18	81.8	607	15	US-10-389-566-1331	Sequence 1331, Ap
108	18	81.8	333	15	US-10-282-122A-64821	Sequence 64821, A	181	18	81.8	607	15	US-10-389-566-2440	Sequence 2440, Ap
109	18	81.8	333	16	US-10-481-265-73	Sequence 73, Appl	182	18	81.8	608	15	US-10-389-566-513	Sequence 513, App
110	18	81.8	334	14	US-10-156-761-8484	Sequence 8484, Ap	183	18	81.8	626	15	US-10-282-122A-69042	Sequence 69042, A
111	18	81.8	334	15	US-10-425-114-44807	Sequence 44807, A	184	18	81.8	627	15	US-10-104-047-3283	Sequence 3283, Ap
112	18	81.8	335	9	US-09-815-242-13346	Sequence 13346, A	185	18	81.8	647	16	US-10-425-115-227439	Sequence 227439,
113	18	81.8	335	15	US-10-282-122A-73914	Sequence 73914, A	186	18	81.8	659	15	US-10-282-132A-58594	Sequence 58594, A
114	18	81.8	335	16	US-10-474-776-654	Sequence 654, App	187	18	81.8	669	16	US-10-437-963-128133	Sequence 128133,
115	18	81.8	335	17	US-10-472-928-1738	Sequence 1738, Ap	188	18	81.8	672	14	US-10-316-161-2	Sequence 2, Appli
116	18	81.8	338	9	US-09-738-626-6959	Sequence 6959, Ap	189	18	81.8	672	15	US-10-094-749-2290	Sequence 2290, Ap
117	18	81.8	343	15	US-10-369-493-16912	Sequence 16912, A	190	18	81.8	697	16	US-10-322-281-729	Sequence 729, App
118	18	81.8	345	15	US-10-369-493-11820	Sequence 11820, A	191	18	81.8	728	15	US-10-369-493-5694	Sequence 5694, Ap
119	18	81.8	347	14	US-10-287-274-346	Sequence 346, App	192	18	81.8	732	15	US-10-369-493-18707	Sequence 18707, A
120	18	81.8	349	15	US-10-389-566-2148	Sequence 2148, Ap	193	18	81.8	766	16	US-10-437-963-178123	Sequence 178123,
121	18	81.8	359	17	US-10-732-923-10163	Sequence 10163, A	194	18	81.8	805	15	US-10-170-385-527	Sequence 527, App
122	18	81.8	366	15	US-10-282-122A-55895	Sequence 55895, A	195	18	81.8	805	16	US-10-393-455-1	Sequence 1, Appli
123	18	81.8	368	15	US-10-369-493-17473	Sequence 17473, A	196	18	81.8	812	16	US-10-128-558-196	Sequence 196, App
124	18	81.8	368	17	US-10-732-923-11847	Sequence 11847, A	197	18	81.8	831	16	US-10-437-963-128135	Sequence 128135,
125	18	81.8	369	15	US-10-424-599-284982	Sequence 284982,	198	18	81.8	836	16	US-10-437-963-122882	Sequence 122882,
126	18	81.8	371	9	US-09-738-626-6691	Sequence 6691, Ap	199	18	81.8	844	15	US-10-225-068-168	Sequence 168, App
127	18	81.8	378	16	US-10-437-963-174114	Sequence 174114,	200	18	81.8	844	17	US-10-374-780A-228	Sequence 228, App
128	18	81.8	387	15	US-10-334-561A-2	Sequence 2, Appli	201	18	81.8	844	17	US-10-225-068-168	Sequence 168, App
129	18	81.8	390	16	US-10-437-963-136516	Sequence 136516,	202	18	81.8	887	15	US-10-282-122A-48500	Sequence 48500, A
130	18	81.8	396	15	US-10-369-493-18706	Sequence 18706, A	203	18	81.8	892	15	US-10-369-493-12293	Sequence 12293, A
131	18	81.8	397	10	US-09-769-744A-152	Sequence 152, App	204	18	81.8	1218	15	US-10-161-493-102	Sequence 102, App
132	18	81.8	397	15	US-10-282-122A-74248	Sequence 74248, A	205	18	81.8	1256	16	US-10-755-889-554	Sequence 554, App
133	18	81.8	397	17	US-10-472-928-4698	Sequence 4698, Ap	206	18	81.8	1526	16	US-10-437-963-117217	Sequence 117217,
134	18	81.8	402	15	US-10-282-122A-72509	Sequence 72509, A	207	18	81.8	2039	14	US-10-192-584-7	Sequence 7, Appli
135	18	81.8	405	15	US-10-389-566-1374	Sequence 1374, Ap	208	18	81.8	2993	9	US-09-738-626-6239	Sequence 6239, Ap
136	18	81.8	405	17	US-10-732-923-1030	Sequence 1030, Ap	209	18	81.8	3913	15	US-10-334-143-45	Sequence 45, Appl
137	18	81.8	412	15	US-10-369-493-12707	Sequence 12707, A	210	18	81.8	4307	15	US-10-369-493-5698	Sequence 5698, Ap
138	18	81.8	415	15	US-10-369-493-7604	Sequence 7604, Ap	211	18	81.8	4307	15	US-10-369-493-5699	Sequence 5699, Ap
139	18	81.8	425	16	US-10-437-963-136517	Sequence 136517,	212	18	81.8	4307	15	US-10-369-493-5700	Sequence 5700, Ap
140	18	81.8	440	15	US-10-442-538-135	Sequence 135, App	213	17	77.3	7	9	US-09-796-264-27	Sequence 27, Appl
141	18	81.8	440	15	US-10-442-538-158	Sequence 158, App	214	17	77.3	7	9	US-09-845-226-27	Sequence 27, Appl
142	18	81.8	447	15	US-10-369-493-4845	Sequence 4845, Ap	215	17	77.3	7	9	US-09-795-903A-27	Sequence 27, Appl
143	18	81.8	451	15	US-10-282-122A-67200	Sequence 67200, A	216	17	77.3	7	16	US-10-820-953-27	Sequence 27, Appl
144	18	81.8	467	14	US-10-325-939-2	Sequence 2, Appli	217	17	77.3	8	9	US-09-796-264-28	Sequence 28, Appl
145	18	81.8	467	17	US-10-732-923-23726	Sequence 23726, A	218	17	77.3	8	9	US-09-845-226-28	Sequence 28, Appl
146	18	81.8	475	13	US-10-083-452-14	Sequence 14, Appl	219	17	77.3	8	9	US-09-795-903A-28	Sequence 28, Appl
147	18	81.8	475	16	US-10-734-510-14	Sequence 14, Appl	220	17	77.3	8	14	US-10-032-818-5	Sequence 5, Appli
148	18	81.8	481	9	US-09-738-626-6986	Sequence 6986, Ap	221	17	77.3	8	16	US-10-281-092-15	Sequence 15, Appl
149	18	81.8	482	14	US-10-156-761-14688	Sequence 14688, A	222	17	77.3	8	16	US-10-281-092-51	Sequence 51, Appl
150	18	81.8	495	15	US-10-369-493-5200	Sequence 5200, Ap	223	17	77.3	8	16	US-10-820-953-28	Sequence 28, Appl
151	18	81.8	495	15	US-10-369-493-5201	Sequence 5201, Ap	224	17	77.3	8	16	US-10-773-754-28	Sequence 28, Appl
152	18	81.8	495	17	US-10-369-493-12162	Sequence 12162, A	225	17	77.3	8	17	US-10-480-954-258	Sequence 258, App
153	18	81.8	495	17	US-10-494-673-14	Sequence 14, Appl	226	17	77.3	8	17	US-10-480-954-259	Sequence 259, App
154	18	81.8	499	13	US-10-108-605-283	Sequence 283, App	227	17	77.3	8	17	US-10-480-954-263	Sequence 263, App
155	18	81.8	501	15	US-10-282-122A-53682	Sequence 53682, A	228	17	77.3	8	17	US-10-480-954-264	Sequence 264, App
156	18	81.8	501	16	US-10-437-963-166123	Sequence 166123,	229	17	77.3	10	10	US-09-908-943A-143	Sequence 143, App
157	18	81.8	507	16	US-10-437-963-150846	Sequence 150846,	230	17	77.3	10	10	US-09-908-943A-144	Sequence 144, App

231 17 77.3 10 16 US-10-801-487-143 Sequence 143, App
232 17 77.3 10 16 US-10-801-487-144 Sequence 144, App
233 17 77.3 10 16 US-10-801-938-143 Sequence 143, App
234 17 77.3 10 16 US-10-801-938-144 Sequence 144, App
235 17 77.3 10 16 US-10-801-509-143 Sequence 143, App
236 17 77.3 10 16 US-10-801-509-144 Sequence 144, App
237 17 77.3 10 16 US-10-801-486-143 Sequence 143, App
238 17 77.3 10 16 US-10-801-486-144 Sequence 144, App
239 17 77.3 10 17 US-10-801-493-143 Sequence 143, App
240 17 77.3 10 17 US-10-801-493-144 Sequence 144, App
241 17 77.3 12 14 US-10-032-818-25 Sequence 25, Appl
242 17 77.3 12 14 US-10-032-818-27 Sequence 27, Appl
243 17 77.3 12 14 US-10-032-818-28 Sequence 28, Appl
244 17 77.3 12 14 US-10-032-818-32 Sequence 32, Appl
245 17 77.3 12 14 US-10-032-818-39 Sequence 39, Appl
246 17 77.3 12 14 US-10-032-818-46 Sequence 46, Appl
247 17 77.3 12 14 US-10-032-818-47 Sequence 47, Appl
248 17 77.3 12 14 US-10-032-818-50 Sequence 50, Appl
249 17 77.3 12 14 US-10-032-818-51 Sequence 51, Appl
250 17 77.3 12 14 US-10-032-818-52 Sequence 52, Appl
251 17 77.3 12 14 US-10-032-818-53 Sequence 53, Appl
252 17 77.3 12 14 US-10-032-818-55 Sequence 55, Appl
253 17 77.3 12 14 US-10-032-818-56 Sequence 56, Appl
254 17 77.3 12 14 US-10-032-818-57 Sequence 57, Appl
255 17 77.3 12 14 US-10-032-818-58 Sequence 58, Appl
256 17 77.3 12 14 US-10-032-818-59 Sequence 59, Appl
257 17 77.3 12 14 US-10-032-818-60 Sequence 60, Appl
258 17 77.3 12 14 US-10-032-818-62 Sequence 62, Appl
259 17 77.3 12 14 US-10-032-818-63 Sequence 63, Appl
260 17 77.3 12 14 US-10-032-818-64 Sequence 64, Appl
261 17 77.3 12 14 US-10-032-818-65 Sequence 65, Appl
262 17 77.3 12 16 US-10-281-092-16 Sequence 16, Appl
263 17 77.3 12 16 US-10-281-092-17 Sequence 17, Appl
264 17 77.3 12 16 US-10-281-092-24 Sequence 24, Appl
265 17 77.3 12 16 US-10-281-092-25 Sequence 25, Appl
266 17 77.3 12 16 US-10-281-092-26 Sequence 26, Appl
267 17 77.3 12 16 US-10-281-092-27 Sequence 27, Appl
268 17 77.3 12 16 US-10-281-092-28 Sequence 28, Appl
269 17 77.3 12 16 US-10-281-092-31 Sequence 31, Appl
270 17 77.3 12 16 US-10-281-092-32 Sequence 32, Appl
271 17 77.3 12 16 US-10-281-092-43 Sequence 43, Appl
272 17 77.3 12 16 US-10-281-092-46 Sequence 46, Appl
273 17 77.3 12 16 US-10-281-092-57 Sequence 57, Appl
274 17 77.3 12 16 US-10-281-092-58 Sequence 58, Appl
275 17 77.3 12 16 US-10-281-092-59 Sequence 59, Appl
276 17 77.3 14 16 US-10-281-092-47 Sequence 47, Appl
277 17 77.3 14 16 US-10-281-092-48 Sequence 48, Appl
278 17 77.3 14 16 US-10-281-092-49 Sequence 49, Appl
279 17 77.3 14 16 US-10-281-092-50 Sequence 50, Appl
280 17 77.3 20 16 US-10-776-013-477 Sequence 477, App
281 17 77.3 23 14 US-10-172-585-24 Sequence 24, Appl
282 17 77.3 23 17 US-10-492-794-19 Sequence 19, Appl
283 17 77.3 32 16 US-10-425-115-286128 Sequence 286128, App
284 17 77.3 32 16 US-10-425-115-301629 Sequence 301629, App
285 17 77.3 36 9 US-09-864-761-41014 Sequence 41014, A
286 17 77.3 43 5 US-10-424-599-183032 Sequence 183032, A
287 17 77.3 44 9 US-09-864-761-38486 Sequence 38486, A
288 17 77.3 48 15 US-10-424-599-225014 Sequence 225014, App
289 17 77.3 54 15 US-10-424-599-226697 Sequence 226697, App
290 17 77.3 58 15 US-10-424-599-215124 Sequence 215124, App
291 17 77.3 59 15 US-10-394-575-47 Sequence 47, Appl
292 17 77.3 62 15 US-10-424-599-264908 Sequence 264908, App
293 17 77.3 63 14 US-10-074-475-273 Sequence 273, App
294 17 77.3 63 16 US-10-767-701-50646 Sequence 50646, A
295 17 77.3 64 15 US-10-424-599-195286 Sequence 195286, App
296 17 77.3 64 16 US-10-425-115-308930 Sequence 308930, App
297 17 77.3 70 15 US-10-424-599-166382 Sequence 166382, App
298 17 77.3 71 14 US-10-252-945-73 Sequence 73, Appl
299 17 77.3 73 9 US-09-925-300-1843 Sequence 1843, App
300 17 77.3 77 16 US-10-425-115-321771 Sequence 321771, App

ALIGNMENTS

RESULT 1

US-10-437-963-200225
; Sequence 200225, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 200225
; LENGTH: 78
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_95715C.1.pep
US-10-437-963-200225

Query Match 86.4%; Score 19; DB 16; Length 78;
Best Local Similarity 57.1%; Pred. No. 6e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VXXAEEF 7
Db 5 VAAAEF 11

RESULT 2

US-10-437-963-161742
; Sequence 161742, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 161742
; LENGTH: 130
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)-(130)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_60899C.1.pep
US-10-437-963-161742

Query Match 86.4%; Score 19; DB 16; Length 130;
Best Local Similarity 57.1%; Pred. No. 1e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
|
Db 36 VSAAEF 42

RESULT 3

US-10-437-963-126883
; Sequence 126883, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 126883
; LENGTH: 146
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_29387C.1.pep
US-10-437-963-126883

Query Match 86.4%; Score 19; DB 16; Length 146;
Best Local Similarity 57.1%; Pred. No. 1.1e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
|
Db 2 VAATAEF 8

RESULT 4

US-10-425-115-361705
; Sequence 361705, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 361705
; LENGTH: 241
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_93051C.1.pep
US-10-425-115-361705

Query Match 86.4%; Score 19; DB 16; Length 241;
Best Local Similarity 57.1%; Pred. No. 1.9e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
|
Db 72 VAAAAEF 78

RESULT 5

US-09-911-781-24
; Sequence 24, Application US/09911781
; Patent No. US20020034788A1
; GENERAL INFORMATION:
; APPLICANT: Craft, David L.
; APPLICANT: Wilson, C. Ron
; APPLICANT: Eirich, Dudley
; APPLICANT: Zhang, Yeyan
; TITLE OF INVENTION: USE OF CYP52A2A PROMOTER TO INCREASE GENE EXPRESSION IN YEAST
; FILE REFERENCE: U0012 OS/OAAP (1010-49)
; CURRENT APPLICATION NUMBER: US/09/911,781
; CURRENT FILING DATE: 2001-07-24
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 24
; LENGTH: 267
; TYPE: PRT
; ORGANISM: Candida tropicalis
US-09-911-781-24

Query Match 86.4%; Score 19; DB 9; Length 267;
Best Local Similarity 57.1%; Pred. No. 2.1e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
|
Db 39 VTTAAEF 45

RESULT 6

US-09-976-800-106
; Sequence 106, Application US/09976800
; Publication No. US2003007795A1
; GENERAL INFORMATION:
; APPLICANT: Wilson, Ron C.
; APPLICANT: Craft, David L.
; APPLICANT: Eirich, Dudley
; APPLICANT: Eshoo, Mark
; APPLICANT: Madduri, Krishna M.
; APPLICANT: Cornett, Cathy A.
; APPLICANT: Brenner, Alfred A.
; APPLICANT: Tang, Maria
; APPLICANT: Loper, John C.
; APPLICANT: Gleeson, Martin
; TITLE OF INVENTION: CYTOCHROME P450 MONOOXYGENASE AND NADPH CYTOCHROME P450 OXIDOREDUCTASE
; TITLE OF INVENTION: GENES AND PROTEINS RELATED TO THE OMEGA HYDROXYLASE COMPLEX OF CYP450
; TITLE OF INVENTION: TROPICALIS AND METHODS RELATING THERETO
; FILE REFERENCE: 1010-16
; CURRENT APPLICATION NUMBER: US/09/976,800
; CURRENT FILING DATE: 2001-10-12
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 106
; LENGTH: 267
; TYPE: PRT
; ORGANISM: CANDIDATROPICALIS
US-09-976-800-106

Query Match 86.4%; Score 19; DB 10; Length 267;
Best Local Similarity 57.1%; Pred. No. 2.1e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
|
Db 39 VTTAAEF 45

RESULT 7

US-10-138-838-106
; Sequence 106, Application US/10138838
; Publication No. US20030049821A1
; GENERAL INFORMATION:


```
; LENGTH: 267
; TYPE: PRT
; ORGANISM: CANDIDATROPICALIS
US-10-138-916-106

Query Match      86.4%; Score 19; DB 14; Length 267;
Best Local Similarity 57.1%; Pred. No. 2.1e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VXXAAEF 7
   |  |||
Db 39 VTTTAEF 45

RESULT 11
US-10-139-296-106
; Sequence 106, Application US/10139296
; Publication No. US20030148486A1
; GENERAL INFORMATION:
; APPLICANT: Wilson, Ron C.
; APPLICANT: Craft, David L.
; APPLICANT: Birch, Dudley
; APPLICANT: Eshoo, Mark
; APPLICANT: Madduri, Krishna M.
; APPLICANT: Cornett, Cathy A.
; APPLICANT: Brenner, Alfred A.
; APPLICANT: Tang, Maria
; APPLICANT: Loper, John C.
; APPLICANT: Gleeson, Martin
; TITLE OF INVENTION: CYTOCHROME P450 MONOOXYGENASE AND NADPH CYTOCHROME
; TITLE OF INVENTION: P450 OXIDOREDUCTASE
; TITLE OF INVENTION: GENES AND PROTEINS RELATED TO THE OMEGA HYDROXYLASE COMPLEX OF
; TITLE OF INVENTION: TROPICALIS AND METHODS RELATING THERETO
; FILE REFERENCE: 1010-16
; CURRENT APPLICATION NUMBER: US/10/139,296
; CURRENT FILING DATE: 2002-05-03
; PRIOR APPLICATION NUMBER: US/09/976,800
; PRIOR FILING DATE: 2001-10-12
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 106
; LENGTH: 267
; TYPE: PRT
; ORGANISM: CANDIDATROPICALIS
US-10-139-296-106

Query Match      86.4%; Score 19; DB 14; Length 267;
Best Local Similarity 57.1%; Pred. No. 2.1e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VXXAAEF 7
   |  |||
Db 39 VTTTAEF 45

RESULT 12
US-10-139-218-106
; Sequence 106, Application US/10139218
; Publication No. US20030153060A1
; GENERAL INFORMATION:
; APPLICANT: Wilson, Ron C.
; APPLICANT: Craft, David L.
; APPLICANT: Birch, Dudley
; APPLICANT: Eshoo, Mark
; APPLICANT: Madduri, Krishna M.
; APPLICANT: Cornett, Cathy A.
; APPLICANT: Brenner, Alfred A.
; APPLICANT: Tang, Maria
; APPLICANT: Loper, John C.
; APPLICANT: Gleeson, Martin
; TITLE OF INVENTION: CYTOCHROME P450 MONOOXYGENASE AND NADPH CYTOCHROME
; TITLE OF INVENTION: P450 OXIDOREDUCTASE
; TITLE OF INVENTION: GENES AND PROTEINS RELATED TO THE OMEGA HYDROXYLASE COMPLEX OF

; TITLE OF INVENTION: TROPICALIS AND METHODS RELATING THERETO
; FILE REFERENCE: 1010-16
; CURRENT APPLICATION NUMBER: US/10/139,218
; CURRENT FILING DATE: 2002-05-03
; PRIOR APPLICATION NUMBER: US/09/976,800
; PRIOR FILING DATE: 2001-10-12
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 106
; LENGTH: 267
; TYPE: PRT
; ORGANISM: CANDIDATROPICALIS
US-10-139-218-106

Query Match      86.4%; Score 19; DB 14; Length 267;
Best Local Similarity 57.1%; Pred. No. 2.1e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VXXAAEF 7
   |  |||
Db 39 VTTTAEF 45

RESULT 13
US-10-400-902-24
; Sequence 24, Application US/10400902
; Publication No. US20030175896A1
; GENERAL INFORMATION:
; APPLICANT: Craft, David L.
; APPLICANT: Wilson, C. Ron
; APPLICANT: Birch, Dudley
; APPLICANT: Zhang, Yeyan
; TITLE OF INVENTION: USE OF CYP52A2A PROMOTER TO INCREASE GENE EXPRESSION IN YEAST
; FILE REFERENCE: U0012 OS/OAAP (1010-49)
; CURRENT APPLICATION NUMBER: US/10/400,902
; CURRENT FILING DATE: 2003-03-31
; PRIOR APPLICATION NUMBER: US/09/911,781
; PRIOR FILING DATE: 2001-07-24
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 24
; LENGTH: 267
; TYPE: PRT
; ORGANISM: Candida tropicalis
US-10-400-902-24

Query Match      86.4%; Score 19; DB 14; Length 267;
Best Local Similarity 57.1%; Pred. No. 2.1e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VXXAAEF 7
   |  |||
Db 39 VTTTAEF 45

RESULT 14
US-10-405-660-106
; Sequence 106, Application US/10405660
; Publication No. US20030186411A1
; GENERAL INFORMATION:
; APPLICANT: Wilson, Ron C.
; APPLICANT: Craft, David L.
; APPLICANT: Birch, Dudley
; APPLICANT: Eshoo, Mark
; APPLICANT: Madduri, Krishna M.
; APPLICANT: Cornett, Cathy A.
; APPLICANT: Brenner, Alfred A.
; APPLICANT: Tang, Maria
; APPLICANT: Loper, John C.
; APPLICANT: Gleeson, Martin
; TITLE OF INVENTION: CYTOCHROME P450 MONOOXYGENASE AND NADPH CYTOCHROME
; TITLE OF INVENTION: GENES AND PROTEINS RELATED TO THE OMEGA HYDROXYLASE COMPLEX OF
; TITLE OF INVENTION: TROPICALIS AND METHODS RELATING THERETO
```


; FILE REFERENCE: 1010-16
; CURRENT APPLICATION NUMBER: US/10/405,660
; CURRENT FILING DATE: 2003-04-02
; PRIOR APPLICATION NUMBER: US/09/976,800
; PRIOR FILING DATE: 2001-10-12
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 106
; LENGTH: 267
; TYPE: PRT
; ORGANISM: CANDIDATROPICALIS
US-10-405-660-106

Query Match 86.4%; Score 19; DB 14; Length 267;
Best Local Similarity 57.1%; Pred. No. 2.1e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VXXAAEF 7
| | | | |
Db 39 VTTAAEF 45

RESULT 15

US-10-138-898-106
; Sequence 106, Application US/1013898
; Publication No. US20030212946A1
; GENERAL INFORMATION:

; APPLICANT: Wilson, Ron C.
; APPLICANT: Craft, David L.
; APPLICANT: Birch, Dudley
; APPLICANT: Eshoo, Mark
; APPLICANT: Madduri, Krishna M.
; APPLICANT: Cornett, Cathy A.
; APPLICANT: Brenner, Alfred A.
; APPLICANT: Tang, Maria
; APPLICANT: Loper, John C.
; APPLICANT: Gleeson, Martin

; TITLE OF INVENTION: CYTOCHROME P450 MONOOXYGENASE AND NADPH CYTOCHROME
; TITLE OF INVENTION: P450 OXIDOREDUCTASE
; TITLE OF INVENTION: GENES AND PROTEINS RELATED TO THE OMEGA HYDROXYLASE COMPLEX OF
; TITLE OF INVENTION: TROPICALIS AND METHODS RELATING THERETO

; FILE REFERENCE: 1010-16
; CURRENT APPLICATION NUMBER: US/10/138,898
; CURRENT FILING DATE: 2002-05-03
; PRIOR APPLICATION NUMBER: US/09/976,800
; PRIOR FILING DATE: 2001-10-12
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 106
; LENGTH: 267
; TYPE: PRT
; ORGANISM: CANDIDATROPICALIS
US-10-138-898-106

Query Match 86.4%; Score 19; DB 15; Length 267;
Best Local Similarity 57.1%; Pred. No. 2.1e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VXXAAEF 7
| | | | |
Db 39 VTTAAEF 45

RESULT 16

US-09-815-242-10242
; Sequence 10242, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:

; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.

; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10242
; LENGTH: 298
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-815-242-10242

Query Match 86.4%; Score 19; DB 9; Length 298;
Best Local Similarity 57.1%; Pred. No. 2.3e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VXXAAEF 7
| | | | |
Db 172 VSTAAEF 178

RESULT 17

US-10-282-122A-56631
; Sequence 56631, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:

; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.

; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625

```

; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 56631
; LENGTH: 298
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-282-122A-56631

Query Match      86.4%; Score 19; DB 15; Length 298;
Best Local Similarity 57.1%; Pred. No. 2.3e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VXXAAEF 7
|-----|
Db 172 VSTTAEF 178

RESULT 18
US-10-425-115-302302
; Sequence 302302, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 302302
; LENGTH: 330
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_38776C.1.pep
US-10-425-115-302302

Query Match      86.4%; Score 19; DB 16; Length 330;
Best Local Similarity 57.1%; Pred. No. 2.6e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VXXAAEF 7
|-----|
Db 107 VASAAEF 113

RESULT 19
US-10-282-122A-72638
; Sequence 72638, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.

```

```

; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 72638
; LENGTH: 338
; TYPE: PRT
; ORGANISM: Salmonella paratyphi A
US-10-282-122A-72638

```

```

Query Match      86.4%; Score 19; DB 15; Length 338;
Best Local Similarity 57.1%; Pred. No. 2.6e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

Qy 1 VXXAAEF 7
|-----|
Db 271 VSSAAEF 277

```

```

RESULT 20
US-10-888-656-2
; Sequence 2, Application US/10888656
; Publication No. US20050106588A1
; GENERAL INFORMATION:
; APPLICANT: Institut Fur Planzenbiochemie
; TITLE OF INVENTION: Reticuline 7-O-methyltransferase
; FILE REFERENCE: B5688
; CURRENT APPLICATION NUMBER: US/10/888,656
; CURRENT FILING DATE: 2004-07-08
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 355
; TYPE: PRT
; ORGANISM: Papaver somniferum
US-10-888-656-2

```

```

Query Match      86.4%; Score 19; DB 17; Length 355;
Best Local Similarity 57.1%; Pred. No. 2.6e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

Qy 1 VXXAAEF 7
|-----|
Db 226 VATAAEF 232

```

```

RESULT 21
US-10-425-114-57008
; Sequence 57008, Application US/10425114

```

; Publication No. US20040034889A1

; GENERAL INFORMATION:

; APPLICANT: Liu, Jingdong

; APPLICANT: Zhou, Yihua

; APPLICANT: Kovalic, David K.

; APPLICANT: Screen, Steven E

; APPLICANT: Tabaska, Jack E

; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

; FILE REFERENCE: 38-21(53313)B

; CURRENT APPLICATION NUMBER: US/10/425,114

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 73128

; SEQ ID NO 57008

; LENGTH: 396

; TYPE: PRT

; ORGANISM: Zea mays

; FEATURE:

; OTHER INFORMATION: Clone ID: UC-2MFLMO17116F06_FLI.pep

US-10-425-114-57008

Query Match 86.4%; Score 19; DB 15; Length 396;

Best Local Similarity 57.1%; Pred. No. 3.1e+03;

Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VXXAAEF 7

Db 226 VTAAAEF 232

RESULT 22

US-10-437-963-128132

; Sequence 128132, Application US/10437963

; Publication No. US2004012343A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei

; APPLICANT: Wu, Wei

; APPLICANT: Boukharov, Andrey A.

; APPLICANT: Barbazuk, Brad

; APPLICANT: Li, Ping

; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With

; FILE REFERENCE: 38-21(53221)B

; CURRENT APPLICATION NUMBER: US/10/437,963

; CURRENT FILING DATE: 2003-05-14

; NUMBER OF SEQ ID NOS: 204966

; SEQ ID NO 128132

; LENGTH: 586

; TYPE: PRT

; ORGANISM: Oryza sativa

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT_MRT4530_30515C.1.pep

US-10-437-963-128132

Query Match 86.4%; Score 19; DB 16; Length 586;

Best Local Similarity 57.1%; Pred. No. 4.6e+03;

Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VXXAAEF 7

Db 370 VSSTAEEF 376

RESULT 23

US-10-425-115-336689

; Sequence 336689, Application US/10425115

; Publication No. US20040214272A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

; FILE REFERENCE: 38-21(53222)B

; CURRENT APPLICATION NUMBER: US/10/425,115

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 369326

; SEQ ID NO 336689

; LENGTH: 747

; TYPE: PRT

; ORGANISM: Zea mays

; FEATURE:

; OTHER INFORMATION: Clone ID: MRT4577_70192C.1.pep

US-10-425-115-336689

Query Match 86.4%; Score 19; DB 16; Length 747;

Best Local Similarity 57.1%; Pred. No. 5.8e+03;

Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VXXAAEF 7

Db 577 VTAAAEF 583

RESULT 24

US-10-425-114-51691

; Sequence 51691, Application US/10425114

; Publication No. US20040034888A1

; GENERAL INFORMATION:

; APPLICANT: Liu, Jingdong

; APPLICANT: Zhou, Yihua

; APPLICANT: Kovalic, David K.

; APPLICANT: Screen, Steven E

; APPLICANT: Tabaska, Jack E

; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

; FILE REFERENCE: 38-21(53313)B

; CURRENT APPLICATION NUMBER: US/10/425,114

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 73128

; SEQ ID NO 51691

; LENGTH: 751

; TYPE: PRT

; ORGANISM: Zea mays

; FEATURE:

; OTHER INFORMATION: Clone ID: 700477144_FLI.pep

US-10-425-114-51691

Query Match 86.4%; Score 19; DB 15; Length 751;

Best Local Similarity 57.1%; Pred. No. 5.9e+03;

Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VXXAAEF 7

Db 581 VTAAAEF 587

RESULT 25

US-10-425-114-62401

; Sequence 62401, Application US/10425114

; Publication No. US20040034888A1

; GENERAL INFORMATION:

; APPLICANT: Liu, Jingdong

; APPLICANT: Zhou, Yihua

; APPLICANT: Kovalic, David K.

; APPLICANT: Screen, Steven E

; APPLICANT: Tabaska, Jack E

; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

; FILE REFERENCE: 38-21(53313)B

; CURRENT APPLICATION NUMBER: US/10/425,114

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 73128

; SEQ ID NO 51691

; LENGTH: 751

; TYPE: PRT

; ORGANISM: Zea mays

; FEATURE:

; OTHER INFORMATION: Clone ID: 700477144_FLI.pep

US-10-425-114-51691

; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 62401
; LENGTH: 751
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFLM017212G02_FLI.pep
US-10-425-114-62401

Query Match 86.4%; Score 19; DB 15; Length 751;
Best Local Similarity 57.1%; Pred. No. 5.9e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
|
|
|
|
Db 581 VTAAAEF 587

RESULT 26
US-10-282-122A-75669
; Sequence 75669, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/233,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 75669
; LENGTH: 1266
; TYPE: PRT
; ORGANISM: Salmonella typhi

US-10-282-122A-75669
Query Match 86.4%; Score 19; DB 15; Length 1266;
Best Local Similarity 57.1%; Pred. No. 9.9e+03;

Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 VXXAAEF 7
|
|
|
|
Db 895 VSSAAEF 901

RESULT 27
US-10-489-740-173
; Sequence 173, Application US/10489740
; Publication No. US20050112574A1
; GENERAL INFORMATION:
; APPLICANT: Bionomics Limited
; TITLE OF INVENTION: P9
; FILE REFERENCE: Angiogenesis_PCT
; CURRENT APPLICATION NUMBER: US/10/489,740
; CURRENT FILING DATE: 2004-03-15
; NUMBER OF SEQ ID NOS: 216
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 173
; LENGTH: 1584
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-489-740-173

Query Match 86.4%; Score 19; DB 17; Length 1584;
Best Local Similarity 57.1%; Pred. No. 1.2e+04;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
|
|
|
|
Db 1549 VSAAAEF 1555

RESULT 28
US-10-369-493-6737
; Sequence 6737, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 6737
; LENGTH: 2214
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-10-369-493-6737

Query Match 86.4%; Score 19; DB 15; Length 2214;
Best Local Similarity 57.1%; Pred. No. 1.7e+04;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
|
|
|
|
Db 1770 VTTSAEF 1776

RESULT 29
US-10-032-818-48
; Sequence 48, Application US/10032818
; Publication No. US20030092629A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Jordan J.N.

; APPLICANT: Koelsch, Gerald
; APPLICANT: Ghosh, Arun K.
; TITLE OF INVENTION: Inhibitors of Memapsin 2 and Use Thereof
; FILE REFERENCE: 2932.1006-007
; CURRENT APPLICATION NUMBER: US/10/032,818
; CURRENT FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: US 60/275,756
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: US 60/258,705
; PRIOR FILING DATE: 2000-12-28
; NUMBER OF SEQ ID NOS: 83
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 48
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic peptide
US-10-032-818-48

Query Match 81.8%; Score 18; DB 14; Length 12;
Best Local Similarity 57.1%; Pred. No. 1.8e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
| |||
Db 6 VNSAAEF 12

RESULT 30
US-10-032-818-49
; Sequence 49, Application US/10032818
; Publication No. US20030092629A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Jordan J.N.
; APPLICANT: Koelsch, Gerald
; TITLE OF INVENTION: Inhibitors of Memapsin 2 and Use Thereof
; FILE REFERENCE: 2932.1006-007
; CURRENT APPLICATION NUMBER: US/10/032,818
; CURRENT FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: US 60/275,756
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: US 60/258,705
; PRIOR FILING DATE: 2000-12-28
; NUMBER OF SEQ ID NOS: 83
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 49
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic peptide
US-10-032-818-49

Query Match 81.8%; Score 18; DB 14; Length 12;
Best Local Similarity 57.1%; Pred. No. 1.8e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
| |||
Db 6 VNTAAEF 12

RESULT 31
US-10-032-818-54
; Sequence 54, Application US/10032818
; Publication No. US20030092629A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Jordan J.N.
; APPLICANT: Koelsch, Gerald
; APPLICANT: Ghosh, Arun K.
; TITLE OF INVENTION: Inhibitors of Memapsin 2 and Use Thereof

; FILE REFERENCE: 2932.1006-007
; CURRENT APPLICATION NUMBER: US/10/032,818
; CURRENT FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: US 60/275,756
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: US 60/258,705
; PRIOR FILING DATE: 2000-12-28
; NUMBER OF SEQ ID NOS: 83
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 54
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic peptide
US-10-032-818-54

Query Match 81.8%; Score 18; DB 14; Length 12;
Best Local Similarity 57.1%; Pred. No. 1.8e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
| |||
Db 6 VNAAEF 12

RESULT 32
US-10-424-599-212707
; Sequence 212707, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 212707
; LENGTH: 49
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_340C.1.pap
US-10-424-599-212707

Query Match 81.8%; Score 18; DB 15; Length 49;
Best Local Similarity 57.1%; Pred. No. 7.3e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
| |||
Db 33 VVSTAEF 39

RESULT 33
US-09-984-429-158
; Sequence 158, Application US/09984429
; Publication No. US20040010132A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 53 Human Secreted Proteins
; FILE REFERENCE: P2018P2
; CURRENT APPLICATION NUMBER: US/09/984,429
; CURRENT FILING DATE: 2001-10-30
; PRIOR APPLICATION NUMBER: 60/244,591
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 09/288,143
; PRIOR FILING DATE: 1999-04-08
; PRIOR APPLICATION NUMBER: PCT/US98/21142

; PRIOR FILING DATE: 1998-10-08
; PRIOR APPLICATION NUMBER: 60/061,463
; PRIOR FILING DATE: 1997-10-09
; PRIOR APPLICATION NUMBER: 60/061,529
; PRIOR FILING DATE: 1997-10-09
; PRIOR APPLICATION NUMBER: 60/071,498
; PRIOR FILING DATE: 1997-10-09
; PRIOR APPLICATION NUMBER: 60/061,527
; PRIOR FILING DATE: 1997-10-09
; PRIOR APPLICATION NUMBER: 60/061,536
; PRIOR FILING DATE: 1997-10-09
; PRIOR APPLICATION NUMBER: 60/061,532
; PRIOR FILING DATE: 1997-10-09
; NUMBER OF SEQ ID NOS: 727
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 158
; LENGTH: 56
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (55)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-984-429-158

Query Match 81.8%; Score 18; DB 11; Length 56;
Best Local Similarity 57.1%; Pred. No. 8.3e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
| |||
Db 36 VAARAEF 42

RESULT 34
US-10-150-111-158
; Sequence 158, Application US/10150111
; Publication No. US20030078386A1
; GENERAL INFORMATION:
; APPLICANT: Rubin et al.
; TITLE OF INVENTION: Secreted Protein HPEAD48
; FILE REFERENCE: P2018P101
; CURRENT APPLICATION NUMBER: US/10/150,111
; CURRENT FILING DATE: 2002-05-20
; PRIOR APPLICATION NUMBER: 09/288,143
; PRIOR FILING DATE: 1999-04-08
; PRIOR APPLICATION NUMBER: PCT/US98/21142
; PRIOR FILING DATE: 1998-10-08
; PRIOR APPLICATION NUMBER: 60/061,463
; PRIOR FILING DATE: 1997-10-09
; PRIOR APPLICATION NUMBER: 60/061,529
; PRIOR FILING DATE: 1997-10-09
; PRIOR APPLICATION NUMBER: 60/071,498
; PRIOR FILING DATE: 1997-10-09
; PRIOR APPLICATION NUMBER: 60/061,527
; PRIOR FILING DATE: 1997-10-09
; PRIOR APPLICATION NUMBER: 60/061,536
; PRIOR FILING DATE: 1997-10-09
; PRIOR APPLICATION NUMBER: 60/061,532
; PRIOR FILING DATE: 1997-10-09
; NUMBER OF SEQ ID NOS: 219
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 158
; LENGTH: 56
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (55)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-150-111-158

Query Match 81.8%; Score 18; DB 14; Length 56;

Best Local Similarity 57.1%; Pred. No. 8.3e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
| |||
Db 36 VAARAEF 42

RESULT 35
US-10-425-115-237245
; Sequence 237245, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 237245
; LENGTH: 60
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_147953C.1.pap
US-10-425-115-237245

Query Match 81.8%; Score 18; DB 16; Length 60;
Best Local Similarity 57.1%; Pred. No. 9e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
| |||
Db 53 VASLAEF 59

RESULT 36
US-10-425-115-243983
; Sequence 243983, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 243983
; LENGTH: 61
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_154096C.1.pap
US-10-425-115-243983

Query Match 81.8%; Score 18; DB 16; Length 61;
Best Local Similarity 57.1%; Pred. No. 9.1e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
| |||
Db 12 VTRAAEF 18

RESULT 37

US-10-425-115-275053
; Sequence 275053, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 275053
; LENGTH: 66
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_182439C.1.pep
US-10-425-115-275053

Query Match 81.8%; Score 18; DB 16; Length 66;
Best Local Similarity 57.1%; Pred. No. 9.9e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
| | | |
DB 59 VAAYAEF 65

RESULT 38
US-10-334-561A-4
; Sequence 4, Application US/10334561A
; Publication No. US20030211520A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel E.
; APPLICANT: Hubert, Rene S.
; APPLICANT: Mitchell, Stephen C.
; TITLE OF INVENTION: NOVEL GENE EXPRESSED IN PROSTATE CANCER
; FILE REFERENCE: 51158203110
; CURRENT APPLICATION NUMBER: US/10/334,561A
; CURRENT FILING DATE: 2002-12-30
; PRIOR APPLICATION NUMBER: 09/410,132
; PRIOR FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: 60/146,584
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: 60/102,572
; PRIOR FILING DATE: 1998-09-30
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 69
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-334-561A-4

Query Match 81.8%; Score 18; DB 15; Length 69;
Best Local Similarity 42.9%; Pred. No. 1e+03;
Matches 3; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
| | | |
DB 1 ISTTAEF 7

RESULT 39
US-10-424-599-228254
; Sequence 228254, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K

; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 228254
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_48141C.1.pep
US-10-424-599-228254

Query Match 81.8%; Score 18; DB 15; Length 86;
Best Local Similarity 57.1%; Pred. No. 1.3e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
| | | |
DB 24 VLSTAEF 30

RESULT 40
US-10-425-114-49255
; Sequence 49255, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 49255
; LENGTH: 91
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700573083_FLI.pep
US-10-425-114-49255

Query Match 81.8%; Score 18; DB 15; Length 91;
Best Local Similarity 57.1%; Pred. No. 1.4e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
| | | |
DB 64 VTSQAEF 70

RESULT 41
US-10-369-493-10066
; Sequence 10066, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493

; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 10066
; LENGTH: 93
; TYPE: PRT
; ORGANISM: magnetite-containing magnetic coccus
US-10-369-493-10066

Query Match 81.8%; Score 18; DB 15; Length 93;
Best Local Similarity 57.1%; Pred. No. 1.4e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
| | | |
Db 22 VRTAAEF 28

RESULT 42

US-10-767-701-53633
; Sequence 53633, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 53633
; LENGTH: 95
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: 13468723.pap
US-10-767-701-53633

Query Match 81.8%; Score 18; DB 16; Length 95;
Best Local Similarity 57.1%; Pred. No. 1.4e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
| | | |
Db 79 VITSAEF 85

RESULT 43

US-10-425-115-255955
; Sequence 255955, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 255955
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_165021C.1.pap
US-10-425-115-255955

Query Match 81.8%; Score 18; DB 16; Length 100;
Best Local Similarity 57.1%; Pred. No. 1.5e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
| | | |
Db 41 VSVSAEF 47

RESULT 44

US-10-425-115-190652
; Sequence 190652, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 190652
; LENGTH: 102
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_105452C.1.pap
US-10-425-115-190652

Query Match 81.8%; Score 18; DB 16; Length 102;
Best Local Similarity 57.1%; Pred. No. 1.5e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
| | | |
Db 83 VSVSAEF 89

RESULT 45

US-10-425-115-326590
; Sequence 326590, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 326590
; LENGTH: 116
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_60920C.1.pap
US-10-425-115-326590

Query Match 81.8%; Score 18; DB 16; Length 116;
Best Local Similarity 57.1%; Pred. No. 1.7e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
| | | |
Db 92 VAANAEF 98

RESULT 46

US-10-424-599-144633
; Sequence 144633, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 144633
; LENGTH: 121
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)-(121)
; OTHER INFORMATION: unsure at all xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_101616C.1.pep
US-10-424-599-144633

Query Match 81.8%; Score 18; DB 15; Length 121;
Best Local Similarity 57.1%; Pred. No. 1.8e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
|
|
|
|
DB 32 VLAAAEF 38

RESULT 47

US-10-437-963-184577
; Sequence 184577, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 184577
; LENGTH: 122
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_81558C.1.pep
US-10-437-963-184577

Query Match 81.8%; Score 18; DB 16; Length 122;
Best Local Similarity 57.1%; Pred. No. 1.8e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
|
|
|
|
DB 91 VISTAEP 97

RESULT 48

US-10-425-115-271612
; Sequence 271612, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 271612
; LENGTH: 122
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_179303C.1.pep
US-10-425-115-271612

Query Match 81.8%; Score 18; DB 16; Length 122;
Best Local Similarity 57.1%; Pred. No. 1.8e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
|
|
|
|
DB 12 VTRAAPF 18

RESULT 49

US-10-425-115-274994
; Sequence 274994, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 274994
; LENGTH: 122
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_182385C.1.pep
US-10-425-115-274994

Query Match 81.8%; Score 18; DB 16; Length 122;
Best Local Similarity 57.1%; Pred. No. 1.8e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
|
|
|
|
DB 12 VTRAAPF 18

RESULT 50

US-10-424-599-156801
; Sequence 156801, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 156801
; LENGTH: 124
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_112613C.1.pep
US-10-424-599-156801

Query Match 81.8%; Score 18; DB 15; Length 124;
Best Local Similarity 57.1%; Pred. No. 1.9e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VXXAAEF 7
| |||
Db 14 VRSAAEF 20

RESULT 51

US-10-767-701-33440
; Sequence 33440, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 33440
; LENGTH: 124
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C32083_1.pep
US-10-767-701-33440

Query Match 81.8%; Score 18; DB 16; Length 124;
Best Local Similarity 57.1%; Pred. No. 1.9e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VXXAAEF 7
| |||
Db 59 VGAAAEF 65

RESULT 52

US-10-425-115-297969
; Sequence 297969, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 297969
; LENGTH: 124
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_48043C.1.pep
US-10-424-599-156801

Query Match 81.8%; Score 18; DB 15; Length 129;
Best Local Similarity 57.1%; Pred. No. 1.9e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VXXAAEF 7
| |||

; OTHER INFORMATION: Clone ID: MRT4577_34819C.1.pep
US-10-425-115-297969

Query Match 81.8%; Score 18; DB 16; Length 124;
Best Local Similarity 57.1%; Pred. No. 1.9e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VXXAAEF 7
| |||
Db 112 VTAAAEF 118

RESULT 53

US-10-425-115-238741
; Sequence 238741, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 238741
; LENGTH: 125
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_14931C.1.pep
US-10-425-115-238741

Query Match 81.8%; Score 18; DB 16; Length 125;
Best Local Similarity 57.1%; Pred. No. 1.9e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VXXAAEF 7
| |||
Db 16 VITAAEF 22

RESULT 54

US-10-424-599-228146
; Sequence 228146, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 228146
; LENGTH: 129
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_48043C.1.pep
US-10-424-599-228146

Query Match 81.8%; Score 18; DB 15; Length 129;
Best Local Similarity 57.1%; Pred. No. 1.9e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VXXAAEF 7
| |||

```
Db      33 VAAAEF 39

RESULT 55
US-10-424-599-274885
; Sequence 274885, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: Zhou Yihua
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 274885
; LENGTH: 137
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_90241C.1.pep
US-10-424-599-274885

Query Match      81.8%; Score 18; DB 15; Length 137;
Best Local Similarity 57.1%; Pred. No. 2.le+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 VXXAAEF 7
      | |||
Db      103 VAAAEF 109

RESULT 56
US-10-370-715B-390
; Sequence 390, Application US/10370715B
; Publication No. US20040258678A1
; GENERAL INFORMATION:
; Patin Docket Preview
; APPLICANT: BODARY, SARAH C.
; APPLICANT: CLARK, HILLARY
; APPLICANT: BRISDELL, HUNTE
; APPLICANT: JACKMAN, JANET
; APPLICANT: SCHOENFELD, JILL R.
; APPLICANT: WILLIAMS, F. MICKEY
; APPLICANT: WOOD, WILLIAM I.
; APPLICANT: WU, THOMAS D.
; TITLE OF INVENTION: Compositions and Methods for the Treatment of Immune
; FILE REFERENCE: P1948R1-US
; CURRENT APPLICATION NUMBER: US/10/370,715B
; CURRENT FILING DATE: 2003-02-21
; NUMBER OF SEQ ID NOS: 742
; SEQ ID NO 390
; LENGTH: 137
; TYPE: PRT
; ORGANISM: Homo sapien
; OTHER INFORMATION: Clone ID: PAT_MRT3847_90241C.1.pep
US-10-370-715B-390

Query Match      81.8%; Score 18; DB 16; Length 137;
Best Local Similarity 57.1%; Pred. No. 2.le+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 VXXAAEF 7
      | |||
Db      7 VITAAEF 13

RESULT 57
US-10-425-114-61704
; Sequence 61704, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 61704
; LENGTH: 141
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3116-011-F10_FLI.pep
US-10-425-114-61704

Query Match      81.8%; Score 18; DB 15; Length 141;
Best Local Similarity 57.1%; Pred. No. 2.le+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 VXXAAEF 7
      | |||
Db      31 VTRADEF 37

RESULT 58
US-10-425-115-274236
; Sequence 274236, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 274236
; LENGTH: 142
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_181693C.1.pep
US-10-425-115-274236

Query Match      81.8%; Score 18; DB 16; Length 142;
Best Local Similarity 57.1%; Pred. No. 2.le+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 VXXAAEF 7
      | |||
Db      107 VAAAEF 113

RESULT 59
US-10-424-599-279916
; Sequence 279916, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
```

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 279916
; LENGTH: 144
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(144)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_94788C.1.pep
US-10-424-599-279916

Query Match 81.8%; Score 18; DB 15; Length 144;
Best Local Similarity 42.9%; Pred. No. 2.2e+03;
Matches 3; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VXXAAEF 7
Db 133 IATAAEF 139

RESULT 60

US-10-437-963-169685
; Sequence 169685, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping

; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 169685
; LENGTH: 149
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(149)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_68081C.1.pep
US-10-437-963-169685

Query Match 81.8%; Score 18; DB 16; Length 149;
Best Local Similarity 57.1%; Pred. No. 2.2e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VXXAAEF 7
Db 117 VGTAAEF 123

RESULT 61

US-10-437-963-198547
; Sequence 198547, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 198547
; LENGTH: 155
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_94196C.1.pep
US-10-437-963-198547

Query Match 81.8%; Score 18; DB 16; Length 155;
Best Local Similarity 57.1%; Pred. No. 2.3e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VXXAAEF 7
Db 14 VTSLAEF 20

RESULT 62

US-10-369-493-16932
; Sequence 16932, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 16932
; LENGTH: 159
; TYPE: PRT
; ORGANISM: Caulobacter crescentus
US-10-369-493-16932

Query Match 81.8%; Score 18; DB 15; Length 159;
Best Local Similarity 57.1%; Pred. No. 2.4e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VXXAAEF 7
Db 142 VAAAEAF 148

RESULT 63

US-10-424-599-234753
; Sequence 234753, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

```
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 234753
; LENGTH: 161
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_54009C.1.pep
US-10-424-599-234753

Query Match      81.8%; Score 18; DB 15; Length 161;
Best Local Similarity 57.1%; Pred. No. 2.4e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 VXXAAEF 7
Db      104 VAARAEF 110

RESULT 64
US-10-424-599-245373
; Sequence 245373, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 245373
; LENGTH: 165
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_63602C.1.pep
US-10-424-599-245373

Query Match      81.8%; Score 18; DB 15; Length 165;
Best Local Similarity 57.1%; Pred. No. 2.5e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 VXXAAEF 7
Db      145 VASAEF 151

RESULT 65
US-10-425-114-64341
; Sequence 64341, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 64341
; LENGTH: 167
; TYPE: PRT
```

```
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3632-058-F7_FLI.pep
US-10-425-114-64341

Query Match      81.8%; Score 18; DB 15; Length 167;
Best Local Similarity 57.1%; Pred. No. 2.5e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 VXXAAEF 7
Db      57 VTRAEF 63

RESULT 66
US-10-156-761-7611
; Sequence 7611, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 7611
; LENGTH: 171
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-7611

Query Match      81.8%; Score 18; DB 14; Length 171;
Best Local Similarity 57.1%; Pred. No. 2.6e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 VXXAAEF 7
Db      43 VASAEF 49

RESULT 67
US-10-264-213-199
; Sequence 199, Application US/10264213
; Publication No. US20040009490A1
; GENERAL INFORMATION:
; APPLICANT: Glenn, Matthew
; APPLICANT: Havukkala, Ilkka J
; APPLICANT: Lubbers, Mark William
; APPLICANT: Dekker, James
; TITLE OF INVENTION: Polynucleotides, materials incorporating
; TITLE OF INVENTION: them, and methods for using them.
; FILE REFERENCE: 11000.1043c3
; CURRENT APPLICATION NUMBER: US/10/264,213
; CURRENT FILING DATE: 2002-10-03
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 199
; LENGTH: 179
; TYPE: PRT
; ORGANISM: Lactobacillus rhamnosus
US-10-264-213-199

Query Match      81.8%; Score 18; DB 15; Length 179;
```

Best Local Similarity 57.1%; Pred. No. 2.7e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VXXAAEF 7
| |||
Db 162 VSTIAEF 168

RESULT 68
US-10-425-115-199164
; Sequence 199164, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 199164
; LENGTH: 183
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(183)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_113213C.1.pep
US-10-425-115-199164

Query Match 81.8%; Score 18; DB 16; Length 183;
Best Local Similarity 57.1%; Pred. No. 2.8e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VXXAAEF 7
| |||
Db 157 VTGAEP 163

RESULT 69
US-10-767-701-40084
; Sequence 40084, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 40084
; LENGTH: 187
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(187)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C43151_1.pep
US-10-767-701-40084

Query Match 81.8%; Score 18; DB 16; Length 187;
Best Local Similarity 57.1%; Pred. No. 2.8e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VXXAAEF 7
| |||
Db 11 VSGAEF 17

RESULT 70
US-10-424-599-162790
; Sequence 162790, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 162790
; LENGTH: 188
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_118017C.1.pep
US-10-424-599-162790

Query Match 81.8%; Score 18; DB 15; Length 188;
Best Local Similarity 57.1%; Pred. No. 2.8e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VXXAAEF 7
| |||
Db 91 VATGAEP 97

RESULT 71
US-10-425-115-241928
; Sequence 241928, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 241928
; LENGTH: 188
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_152217C.1.pep
US-10-425-115-241928

Query Match 81.8%; Score 18; DB 16; Length 188;
Best Local Similarity 57.1%; Pred. No. 2.8e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VXXAAEF 7
| |||
Db 52 VAVAEF 58

RESULT 72
US-10-156-761-9937
; Sequence 9937, Application US/10156761

; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 9937
; LENGTH: 194
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-9937

Query Match 81.8%; Score 18; DB 14; Length 194;
Best Local Similarity 57.1%; Pred. No. 2.9e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VXXAAEF 7
Db 58 VSITAEF 64

RESULT 73
US-10-425-115-271613
; Sequence 271613, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 271613
; LENGTH: 199
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_179304C.1.pep
US-10-425-115-271613

Query Match 81.8%; Score 18; DB 16; Length 199;
Best Local Similarity 57.1%; Pred. No. 3e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VXXAAEF 7
Db 89 VTRADEF 95

RESULT 74
US-10-437-963-165533
; Sequence 165533, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 165533
; LENGTH: 201
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(201)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_6432C.1.pep
US-10-437-963-165533

Query Match 81.8%; Score 18; DB 16; Length 201;
Best Local Similarity 57.1%; Pred. No. 3e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VXXAAEF 7
Db 178 VAAGAEF 184

RESULT 75
US-10-437-963-121186
; Sequence 121186, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 121186
; LENGTH: 209
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_24236C.1.pep
US-10-437-963-121186

Query Match 81.8%; Score 18; DB 16; Length 209;
Best Local Similarity 57.1%; Pred. No. 3.2e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VXXAAEF 7
Db 91 VTGADEF 97

RESULT 76
US-10-424-599-230315
; Sequence 230315, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:

; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 230315
; LENGTH: 212
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(212)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_49999C.1.pep
US-10-424-599-230315

Query Match 81.8%; Score 18; DB 15; Length 212;
Best Local Similarity 57.1%; Pred. No. 3.2e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
| | | |
Db 60 VSSLAEF 66

RESULT 77
US-10-424-599-149058
; Sequence 149058, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 149058
; LENGTH: 213
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_105622C.1.pep
US-10-424-599-149058

Query Match 81.8%; Score 18; DB 15; Length 213;
Best Local Similarity 57.1%; Pred. No. 3.2e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
| | | |
Db 56 VSKAAEF 62

RESULT 78
US-10-424-599-149059
; Sequence 149059, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 149059
; LENGTH: 213
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_105623C.1.pep
US-10-424-599-149059

Query Match 81.8%; Score 18; DB 15; Length 213;
Best Local Similarity 57.1%; Pred. No. 3.2e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
| | | |
Db 56 VSKAAEF 62

RESULT 79
US-10-425-115-199685
; Sequence 199685, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 199685
; LENGTH: 220
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_113692C.1.pep
US-10-425-115-199685

Query Match 81.8%; Score 18; DB 16; Length 220;
Best Local Similarity 57.1%; Pred. No. 3.3e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
| | | |
Db 155 VGAARF 161

RESULT 80
US-10-389-566-1581
; Sequence 1581, Application US/10389566
; Publication No. US20040025202A1
; GENERAL INFORMATION:
; APPLICANT: Monsanto Technology, LLC
; APPLICANT: Laurie, Cathy C
; TITLE OF INVENTION: Nucleic Acid Molecules Associated with Oil in Plants
; FILE REFERENCE: 38-77(52900)D
; CURRENT APPLICATION NUMBER: US/10/389,566
; CURRENT FILING DATE: 2003-03-31
; PRIOR APPLICATION NUMBER: US 60/365,301
; PRIOR FILING DATE: 2002-03-15
; PRIOR APPLICATION NUMBER: US 60/391,786
; PRIOR FILING DATE: 2002-06-25
; PRIOR APPLICATION NUMBER: US 60/392,018
; PRIOR FILING DATE: 2002-06-26
; NUMBER OF SEQ ID NOS: 2459
; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 1581
; LENGTH: 228
; TYPE: PRT
; ORGANISM: Vicia faba
US-10-389-566-1581

Query Match 81.8%; Score 18; DB 15; Length 228;
Best Local Similarity 57.1%; Pred. No. 3.4e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
DB 41 VEAAEF 47

RESULT 81

US-10-389-566-1597
; Sequence 1597, Application US/10389566
; Publication No. US20040025202A1
; GENERAL INFORMATION:
; APPLICANT: Monsanto Technology, LLC
; TITLE OF INVENTION: Nucleic Acid Molecules Associated with Oil in Plants
; FILE REFERENCE: 38-77(52900)D
; CURRENT FILING DATE: 2003-03-31
; PRIOR APPLICATION NUMBER: US/10/389,566
; PRIOR FILING DATE: 2002-03-15
; PRIOR APPLICATION NUMBER: US 60/365,301
; PRIOR FILING DATE: 2002-03-15
; PRIOR APPLICATION NUMBER: US 60/391,786
; PRIOR FILING DATE: 2002-06-25
; PRIOR APPLICATION NUMBER: US 60/392,018
; PRIOR FILING DATE: 2002-06-26
; NUMBER OF SEQ ID NOS: 2459
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 1597
; LENGTH: 228
; TYPE: PRT
; ORGANISM: Pisum sativum
US-10-389-566-1597

Query Match 81.8%; Score 18; DB 15; Length 228;
Best Local Similarity 57.1%; Pred. No. 3.4e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
DB 41 VEAAEF 47

RESULT 82

US-10-424-599-247828
; Sequence 247828, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT FILING DATE: 2003-04-28
; CURRENT APPLICATION NUMBER: US/10/424,599
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 247828
; LENGTH: 234
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_65819C.1.p
US-10-424-599-247828

Query Match 81.8%; Score 18; DB 15; Length 234;

Best Local Similarity 57.1%; Pred. No. 3.5e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
DB 151 VSSLAEF 157

RESULT 83

US-10-893-671-98
; Sequence 98, Application US/10893671
; Publication No. US20050064527A1
; GENERAL INFORMATION:
; APPLICANT: Levy, Stuart, et. al.
; FILE REFERENCE: PKZ-043
; CURRENT APPLICATION NUMBER: US/10/893,671
; CURRENT FILING DATE: 2004-07-15
; PRIOR APPLICATION NUMBER: US/09/801,563
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 60/188,362
; PRIOR FILING DATE: 2000-03-10
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 98
; LENGTH: 240
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-893-671-98

Query Match 81.8%; Score 18; DB 17; Length 240;
Best Local Similarity 57.1%; Pred. No. 3.6e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
DB 70 VAQAAEF 76

RESULT 84

US-10-424-599-262409
; Sequence 262409, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT FILING DATE: 2003-04-28
; CURRENT APPLICATION NUMBER: US/10/424,599
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 262409
; LENGTH: 241
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_78979C.1.p
US-10-424-599-262409

Query Match 81.8%; Score 18; DB 15; Length 241;
Best Local Similarity 57.1%; Pred. No. 3.6e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
DB 70 VSLAAEF 76

RESULT 85

US-10-437-963-115265

; Sequence 115265, Application US/10437963
; Publication No. US20040123342A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 115265
; LENGTH: 252
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_1887C.1.pep
US-10-437-963-115265

Query Match 81.8%; Score 18; DB 16; Length 252;
Best Local Similarity 57.1%; Pred. No. 3.8e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VXXAAEF 7
| |||
Db 199 VTWSAEF 205

RESULT 86
US-10-762-107-46
; Sequence 46, Application US/10762107
; Publication No. US20050043297A1
; GENERAL INFORMATION:
; APPLICANT: ECOPIA BIOSCIENCES INC.
; APPLICANT: Farnet, Chris
; APPLICANT: McAlpine, James
; APPLICANT: Zazopoulos, Emmanuel
; APPLICANT: Bachmann, Brian
; APPLICANT: Pirae, Mahmood
; TITLE OF INVENTION: FARNESYL DIBENZODIAZEPINONE, PROCESSES FOR ITS PRODUCTION AND ITS
; FILE REFERENCE: 3005-5US
; CURRENT APPLICATION NUMBER: US/10/762,107
; CURRENT FILING DATE: 2004-01-20
; PRIOR APPLICATION NUMBER: USN 60/441,126
; PRIOR FILING DATE: 2003-01-21
; PRIOR APPLICATION NUMBER: USN 60/492,997
; PRIOR FILING DATE: 2003-08-07
; PRIOR APPLICATION NUMBER: USN 60/518,286
; PRIOR FILING DATE: 2003-11-10
; NUMBER OF SEQ ID NOS: 89
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 46
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Micromonospora sp. strain 046-ECO11
US-10-762-107-46

Query Match 81.8%; Score 18; DB 17; Length 253;
Best Local Similarity 57.1%; Pred. No. 3.8e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VXXAAEF 7
| |||
Db 123 VAVAAEF 129

RESULT 87
US-10-424-599-274705
; Sequence 274705, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 274705
; LENGTH: 256
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_9007C.1.pep
US-10-424-599-274705

Query Match 81.8%; Score 18; DB 15; Length 256;
Best Local Similarity 57.1%; Pred. No. 3.9e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VXXAAEF 7
| |||
Db 22 VFTTAEF 28

RESULT 88
US-09-738-626-5807
; Sequence 5807, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 5807
; LENGTH: 258
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-5807

Query Match 81.8%; Score 18; DB 9; Length 258;
Best Local Similarity 57.1%; Pred. No. 3.9e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VXXAAEF 7
| |||
Db 46 VATGAEF 52

```
RESULT 89
US-10-781-014-600
; Sequence 600, Application US/10781014
; Publication No. US20040180408A1
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Krogger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberhauer, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
; TITLE OF INVENTION: INVOLVED IN CARBON METABOLISM AND ENERGY
; FILE OF INVENTION: PRODUCTION
; FILE REFERENCE: BGI-126CPCN
; CURRENT APPLICATION NUMBER: US/10/781,014
; CURRENT FILING DATE: 2004-02-17
; PRIOR APPLICATION NUMBER: US 09/602,740
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 60/141,031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: 60/143,208
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 60/151,572
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: DE 19931412.8
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931413.6
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931419.5
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931420.9
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931424.1
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931428.4
; PRIOR FILING DATE: 1999-07-08
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 784
; SEQ ID NO 600
; LENGTH: 258
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-10-781-014-600

Query Match      81.8%; Score 18; DB 16; Length 258;
Best Local Similarity 57.1%; Pred. No. 3.9e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 VXXAAEF 7
DB      46 VATGAEF 52
      |||||

RESULT 90
US-10-437-963-114362
; Sequence 114362, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazov, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
```

```
; SEQ ID NO 114362
; LENGTH: 259
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_18060C.1.pep
US-10-437-963-114362

Query Match      81.8%; Score 18; DB 16; Length 259;
Best Local Similarity 57.1%; Pred. No. 3.9e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 VXXAAEF 7
DB      208 VSVSAEF 214
      |||||

RESULT 91
US-10-282-122A-65087
; Sequence 65087, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 65087
; LENGTH: 268
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-282-122A-65087

Query Match      81.8%; Score 18; DB 15; Length 268;
Best Local Similarity 57.1%; Pred. No. 4.1e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 VXXAAEF 7
      |||||
```

```
Db          41 VASNAEF 47

RESULT 92
US-10-282-122A-66020
; Sequence 66020, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 66020
; LENGTH: 268
; TYPE: PRT
; ORGANISM: Neisseria meningitidis
US-10-282-122A-66020

Query Match          81.8%; Score 18; DB 15; Length 268;
Best Local Similarity 57.1%; Pred. No. 4.1e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy          1 VXXXXAEF 7
           |  |||
Db          41 VASNAEF 47

RESULT 93
US-10-275-026A-170
; Sequence 170, Application US/10275026A
; Publication No. US2004008770A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Christoph
; TITLE OF INVENTION: Virulence Genes, Proteins, and Their Use
; FILE REFERENCE: GJE-6436
; CURRENT APPLICATION NUMBER: US/10/275,026A
; CURRENT FILING DATE: 2003-10-31
; PRIOR APPLICATION NUMBER: PCT/GS01/02003

Db          41 VASNAEF 47

Query Match          81.8%; Score 18; DB 15; Length 268;
Best Local Similarity 57.1%; Pred. No. 4.1e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy          1 VXXXXAEF 7
           |  |||
Db          41 VASNAEF 47

Query Match          81.8%; Score 18; DB 15; Length 271;
Best Local Similarity 57.1%; Pred. No. 4.1e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy          1 VXXXXAEF 7
           |  |||
Db          106 VGSSAEF 112

RESULT 95
US-10-282-122A-72181
; Sequence 72181, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
```

; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 72181
; LENGTH: 288
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-10-282-122A-72181

Query Match 81.8%; Score 18; DB 15; Length 288;
Best Local Similarity 57.1%; Pred. No. 4.4e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
| | | |
Db 116 VTGTAEF 122

RESULT 96
US-10-767-701-34443
; Sequence 34443, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 34443
; LENGTH: 293
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(293)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C48605_1.pep
US-10-767-701-34443

Query Match 81.8%; Score 18; DB 16; Length 293;
Best Local Similarity 57.1%; Pred. No. 4.4e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
| | | |
Db 103 VAHAAEF 109

RESULT 97

US-10-425-114-38327
; Sequence 38327, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 38327
; LENGTH: 296
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3242-388-G1_FLI.pep
US-10-425-114-38327

Query Match 81.8%; Score 18; DB 15; Length 296;
Best Local Similarity 57.1%; Pred. No. 4.5e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
| | | |
Db 182 VSVAAEF 188

RESULT 98

US-10-424-599-265888
; Sequence 265888, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 265888
; LENGTH: 301
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_82117C.1.pep
US-10-424-599-265888

Query Match 81.8%; Score 18; DB 15; Length 301;
Best Local Similarity 57.1%; Pred. No. 4.6e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
| | | |
Db 187 VSVAAEF 193

RESULT 99

US-10-369-493-9623
; Sequence 9623, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:

```
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 9623
; LENGTH: 312
; TYPE: PRT
; ORGANISM: Desulfitobacterium hafnense
US-10-369-493-9623

Query Match      81.8%; Score 18; DB 15; Length 312;
Best Local Similarity 57.1%; Pred. No. 4.7e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 VXXAAEF 7
Db      181 VNAAEF 187

RESULT 100
US-10-282-122A-52547
; Sequence 52547, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 52547
```

```
; LENGTH: 315
; TYPE: PRT
; ORGANISM: Clostridium botulinum
US-10-282-122A-52547

Query Match      81.8%; Score 18; DB 15; Length 315;
Best Local Similarity 57.1%; Pred. No. 4.8e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 VXXAAEF 7
Db      115 VATLAEF 121

Search completed: June 13, 2005, 14:18:24
Job time : 112 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2005, 13:36:32 ; Search time 24 Seconds
(without alignments)
28.063 Million cell updates/sec

Title: 09730329-59ED
Perfect score: 22
Sequence: 1 VXXXAEEF 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 300 summaries

Database :

PIR_79:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	19	86.4	240	2 D82288	NADPH-flavin oxido
2	19	86.4	298	2 F85885	probable regulator
3	19	86.4	298	2 C65017	hypothetical prote
4	19	86.4	298	2 C91041	probable regulator
5	19	86.4	385	2 A86165	protein F15K9.6 (i
6	19	86.4	484	2 T31224	traH protein homol
7	19	86.4	1266	2 AF0911	probable exported
8	19	86.4	1460	2 T00095	hypothetical prote
9	19	86.4	1946	2 AE1449	hypothetical prote
10	19	86.4	2214	2 T16305	hypothetical prote
11	18	81.8	76	2 A54252	omega-agatoxin III
12	18	81.8	76	2 D54252	omega-agatoxin III
13	18	81.8	76	2 B54252	omega-agatoxin III
14	18	81.8	76	2 C54252	omega-agatoxin III
15	18	81.8	76	2 A42335	omega-agatoxin III
16	18	81.8	92	2 A34520	calcium-binding pr
17	18	81.8	121	2 S48420	probable membrane
18	18	81.8	146	2 T06471	core protein - gar
19	18	81.8	147	2 H84341	hypothetical prote
20	18	81.8	148	2 A86878	non-heme iron-bind
21	18	81.8	159	2 C87486	hypothetical prote
22	18	81.8	190	2 C87486	calretinin - mouse
23	18	81.8	194	2 T34819	hypothetical prote
24	18	81.8	211	2 S26078	ribosomal protein
25	18	81.8	212	2 I50543	aryl hydrocarbon r
26	18	81.8	212	2 H81344	hypothetical prote
27	18	81.8	213	2 E97174	phosphoserine phos
28	18	81.8	228	2 S65425	pyruvate decarboxy
29	18	81.8	228	2 S65426	pyruvate decarboxy

30	18	81.8	240	2 AG0605	oxygen-insensitive
31	18	81.8	240	2 I80318	drug activity modu
32	18	81.8	240	2 C90745	modulator of drug
33	18	81.8	240	2 G85595	modulator of drug
34	18	81.8	247	2 PQ0662	outer capsid spike
35	18	81.8	247	2 PQ0652	outer capsid spike
36	18	81.8	247	2 PQ0661	outer capsid spike
37	18	81.8	247	2 PQ0655	outer capsid spike
38	18	81.8	247	2 PQ0651	outer capsid spike
39	18	81.8	247	2 PQ0656	outer capsid spike
40	18	81.8	247	2 PQ0660	outer capsid spike
41	18	81.8	247	2 PQ0659	outer capsid spike
42	18	81.8	247	2 PQ0658	outer capsid spike
43	18	81.8	247	2 PQ0653	outer capsid spike
44	18	81.8	247	2 PQ0657	outer capsid spike
45	18	81.8	247	2 S61221	outer capsid spike
46	18	81.8	247	2 S61224	outer capsid spike
47	18	81.8	247	2 S61222	outer capsid spike
48	18	81.8	249	2 S74373	heat shock protein
49	18	81.8	254	2 S34724	probable oxidoredu
50	18	81.8	260	2 I40886	glycine hydroxymet
51	18	81.8	260	2 F69337	thiamin biosynthes
52	18	81.8	261	2 S50699	orotidine-5'-phosp
53	18	81.8	268	2 S19184	hypothetical prote
54	18	81.8	268	2 H81806	histidine-binding
55	18	81.8	268	2 B81063	amino acid ABC tra
56	18	81.8	271	1 A60253	calretinin - human
57	18	81.8	271	1 S25006	calretinin - rat
58	18	81.8	286	2 D83914	chloramphenicol re
59	18	81.8	293	2 H83449	probable transcrip
60	18	81.8	296	2 F84633	probable carbonyl
61	18	81.8	301	1 H69312	conserved hypother
62	18	81.8	316	2 T50027	annexin-like prote
63	18	81.8	317	2 A84304	hypothetical prote
64	18	81.8	320	2 B72656	hypothetical prote
65	18	81.8	321	2 C70653	probable prephenat
66	18	81.8	322	2 F86918	probable prephenat
67	18	81.8	327	2 T35322	hypothetical prote
68	18	81.8	333	1 JCS166	ketol-acid reducto
69	18	81.8	333	2 A69059	ketol-acid reducto
70	18	81.8	333	2 H87120	ketol-acid reducto
71	18	81.8	333	2 D70855	probable ilvc prot
72	18	81.8	335	2 F95103	6-phosphofructokin
73	18	81.8	335	2 D97971	6-phosphofructokin
74	18	81.8	338	1 C48648	ketol-acid reducto
75	18	81.8	342	2 C69313	conserved hypother
76	18	81.8	343	2 T45415	ketol-acid reducto
77	18	81.8	343	2 B87463	hypothetical prote
78	18	81.8	346	2 C82156	conserved hypother
79	18	81.8	347	2 F65131	hypothetical 39.3
80	18	81.8	347	2 F91156	probable transport
81	18	81.8	347	2 C86002	probable transport
82	18	81.8	363	2 AH1427	phosphoserine amin
83	18	81.8	363	2 AF1801	phosphoserine amin
84	18	81.8	365	2 H83369	hypothetical prote
85	18	81.8	368	2 D84140	citrate synthase I
86	18	81.8	371	2 T20093	hypothetical prote
87	18	81.8	391	2 B72283	maltose ABC transp
88	18	81.8	392	2 G97330	probable aminopept
89	18	81.8	393	2 E95261	serine proteinase
90	18	81.8	396	2 C84394	argininosuccinate
91	18	81.8	397	2 F98127	serine proteinase
92	18	81.8	405	2 G84200	cytochrome P450 (i
93	18	81.8	405	2 S65471	pyruvate decarboxy
94	18	81.8	412	2 C88351	phosphoprotein pho
95	18	81.8	417	2 S57820	pyruvate decarboxy
96	18	81.8	420	2 D95972	probable sugar upt
97	18	81.8	438	2 S56631	virion protein kin
98	18	81.8	443	2 S65685	protein phosphatas
99	18	81.8	443	2 B38351	phosphoprotein pho
100	18	81.8	447	2 A38351	phosphoprotein pho
101	18	81.8	447	2 A41805	phosphoprotein pho
102	18	81.8	463	2 T50163	protein phosphatas

103	18	81.8	468	2	JC5417	phosphoprotein pho	176	17	77.3	109	2	C90768	probable minor tai
104	18	81.8	468	2	S65951	[phosphorylase] ph	177	17	77.3	109	2	C85742	probable tail comp
105	18	81.8	473	1	S26191	nitrogenase (EC 1.	178	17	77.3	110	2	T45387	hypothetical prote
106	18	81.8	481	2	S11379	threonine synthase	179	17	77.3	121	2	I40652	RNase P protein co
107	18	81.8	487	2	C95279	GabD3 succinate-se	180	17	77.3	121	4	S58332	hypothetical prote
108	18	81.8	491	2	T38448	probable alpha-amy	181	17	77.3	133	2	T24099	hypothetical prote
109	18	81.8	495	2	T21422	hypothetical prote	182	17	77.3	134	2	T35078	conserved hypothet
110	18	81.8	497	2	T23563	hypothetical prote	183	17	77.3	135	2	H81290	hypothetical prote
111	18	81.8	499	2	A45778	phosphoprotein pho	184	17	77.3	140	2	C75092	hypothetical prote
112	18	81.8	500	2	B86312	type 2A protein se	185	17	77.3	143	2	AF2781	hypothetical prote
113	18	81.8	508	2	S57751	protein phosphatas	186	17	77.3	146	2	C64636	hypothetical prote
114	18	81.8	510	2	A56142	phosphoglycerate m	187	17	77.3	153	2	A69271	hypothetical prote
115	18	81.8	511	2	T17298	hypothetical prote	188	17	77.3	157	2	S43227	homotetic protein P
116	18	81.8	512	2	G96555	55 kDa B regulator	189	17	77.3	161	2	E87455	conserved hypothet
117	18	81.8	513	1	BMH06	bone morphogenetic	190	17	77.3	165	2	S60665	dihydrofolate redu
118	18	81.8	513	2	S55989	protein phosphatas	191	17	77.3	168	2	AB0622	probable bacteriop
119	18	81.8	514	2	T10559	hypothetical prote	192	17	77.3	170	1	HMECBM	M-agglutinin precu
120	18	81.8	526	2	A41698	cell division cont	193	17	77.3	171	2	T37090	probable secreted
121	18	81.8	544	2	T27444	hypothetical prote	194	17	77.3	175	2	C69398	conserved hypothet
122	18	81.8	584	2	B90281	hypothetical prote	195	17	77.3	176	2	H95272	hypothetical prote
123	18	81.8	586	2	T08293	hypothetical prote	196	17	77.3	182	2	T51128	isopentenyl diphos
124	18	81.8	602	2	AE1261	oligopeptidase hom	197	17	77.3	185	2	T33039	hypothetical prote
125	18	81.8	612	2	T38714	hypothetical prote	198	17	77.3	185	2	A39658	polyposis coli reg
126	18	81.8	612	2	S62930	hypothetical prote	199	17	77.3	185	2	JC4667	TB2/Dpl protein ho
127	18	81.8	612	2	S62956	hypothetical prote	200	17	77.3	185	2	A83267	probable transcrip
128	18	81.8	612	2	G83307	hypothetical prote	201	17	77.3	187	2	T32826	hypothetical prote
129	18	81.8	642	2	D64491	hypothetical prote	202	17	77.3	190	2	F84366	hypothetical prote
130	18	81.8	659	2	A44139	ribonuclease T2 ho	203	17	77.3	191	2	T07188	reverse transcript
131	18	81.8	662	2	S61193	probable membrane	204	17	77.3	194	1	RPECRS	resolvase - Escher
132	18	81.8	697	2	A86402	protein T22C5.17 (205	17	77.3	196	2	D64909	probable resolvase
133	18	81.8	710	2	T47610	hypothetical prote	206	17	77.3	196	2	A64888	probable resolvase
134	18	81.8	711	2	C49219	toxin apxIII secre	207	17	77.3	198	2	E75496	conserved hypothet
135	18	81.8	718	2	T40823	probable para-amin	208	17	77.3	203	2	T02868	probable GTP-bindi
136	18	81.8	725	2	T08989	hypothetical prote	209	17	77.3	203	2	T41226	probable ser-thr p
137	18	81.8	728	2	S43606	R07E5.8 protein (c	210	17	77.3	205	2	F72422	2-dehydro-3-deoxy-
138	18	81.8	732	2	F84394	helicase [imported	211	17	77.3	213	2	T44818	NADH2 dehydrogenas
139	18	81.8	732	1	JQ1638	outer layer protei	212	17	77.3	213	2	A84248	transcription regu
140	18	81.8	775	1	JQ1639	outer layer protei	213	17	77.3	213	2	H69251	hypothetical prote
141	18	81.8	792	2	B82752	penicillin binding	214	17	77.3	216	2	E97240	transcription regu
142	18	81.8	812	1	A36477	fibroblast growth	215	17	77.3	219	2	A71879	hypothetical prote
143	18	81.8	814	1	A39732	fibroblast growth	216	17	77.3	221	2	T19468	hypothetical prote
144	18	81.8	824	2	G87338	TonB-dependent rec	217	17	77.3	223	2	H70529	hypothetical prote
145	18	81.8	831	2	G82815	penicillin binding	218	17	77.3	226	2	AG2327	deoxyribose-phosph
146	18	81.8	834	2	D57282	ankyrin-related pr	219	17	77.3	226	2	AC1757	hypothetical prote
147	18	81.8	844	2	T00529	hypothetical prote	220	17	77.3	226	2	AE1650	hypothetical prote
148	18	81.8	862	2	B88594	protein Y49A6B.11	221	17	77.3	226	2	AF1589	hypothetical prote
149	18	81.8	905	2	B71562	probable oxoglutar	222	17	77.3	228	2	A12486	hypothetical prote
150	18	81.8	947	2	T26314	hypothetical prote	223	17	77.3	236	2	D64855	probable transcrip
151	18	81.8	958	2	C87504	hypothetical prote	224	17	77.3	236	2	A90815	hypothetical prote
152	18	81.8	1040	2	S34839	pol protein - frui	225	17	77.3	236	2	E85674	hypothetical prote
153	18	81.8	1056	2	T02930	lysine-ketoglutar	226	17	77.3	244	2	S23099	lectin I, anti-H(O
154	18	81.8	1251	2	A57293	latent transformin	227	17	77.3	246	2	S20940	DNA-binding protei
155	18	81.8	1515	2	B46602	glutamate synthase	228	17	77.3	248	2	S61960	hypothetical prote
156	18	81.8	2344	2	T41590	probable sensor-li	229	17	77.3	250	2	G87327	ABC transporter, A
157	18	81.8	4307	2	T20721	hypothetical prote	230	17	77.3	251	1	B64456	bioC protein homol
158	18	81.8	4377	2	A55575	ankyrin 3, long sp	231	17	77.3	254	2	G65014	hypothetical 29.2K
159	18	81.8	5170	2	T15348	hypothetical prote	232	17	77.3	254	2	G69495	transcription regu
160	17	77.3	59	2	C72534	hypothetical prote	233	17	77.3	257	2	T08304	sojD protein - Hal
161	17	77.3	61	2	F36853	A50L protein - var	234	17	77.3	257	2	E82278	RNA methyltransfer
162	17	77.3	61	2	T72170	A54L protein - var	235	17	77.3	257	2	D82174	conserved hypothet
163	17	77.3	64	2	T34792	probable transcrip	236	17	77.3	257	2	S76930	hypothetical prote
164	17	77.3	70	2	G82833	hypothetical prote	237	17	77.3	258	2	T36359	hypothetical prote
165	17	77.3	76	2	FQ0113	hypothetical prote	238	17	77.3	259	2	F95258	ribosomal protein
166	17	77.3	77	2	B82407	hypothetical prote	239	17	77.3	259	2	A99524	30S ribosomal prot
167	17	77.3	90	2	A70600	hypothetical prote	240	17	77.3	267	2	B99188	thiazole biosynthe
168	17	77.3	95	2	T45170	hypothetical prote	241	17	77.3	270	2	G72390	dihydroorotate deh
169	17	77.3	109	1	TLBPMI	minor tail protein	242	17	77.3	271	2	D70925	probable translati
170	17	77.3	109	2	S27053	thioredoxin - Emer	243	17	77.3	272	2	T01148	probable acetone-c
171	17	77.3	109	2	C85817	hypothetical prote	244	17	77.3	275	2	G75130	translation initia
172	17	77.3	109	2	D90969	probable minor tai	245	17	77.3	275	2	C83482	2-phosphonoacetal
173	17	77.3	109	2	E90899	probable minor tai	246	17	77.3	275	2	D71087	probable translati
174	17	77.3	109	2	D90997	probable minor tai	247	17	77.3	276	2	G87108	elongation factor
175	17	77.3	109	2	D90834	minor tail protein	248	17	77.3	278	2	D75035	met-10+ related pr

249 17 77.3 279 2 T41124 single-stranded DN
 250 17 77.3 281 2 T01445 hypothetical prote
 251 17 77.3 283 2 A44814 endoflagellar prot
 252 17 77.3 283 2 H97087 methyl-accepting c
 253 17 77.3 288 2 AD0884 probable transfera
 254 17 77.3 288 2 S18990 probable acyl-CoA
 255 17 77.3 290 2 G71304 probable translati
 256 17 77.3 295 2 C64123 geranyltransf
 257 17 77.3 295 2 A23750 30k actin-bundling
 258 17 77.3 296 1 S55795 arginase (EC 3.5.3
 259 17 77.3 296 2 H83456 probable oxidoredu
 260 17 77.3 296 2 E83285 probable transcrip
 261 17 77.3 296 2 T47914 hypothetical prote
 262 17 77.3 299 1 S68863 arginase (EC 3.5.3
 263 17 77.3 299 2 A41900 cyn operon regulat
 264 17 77.3 299 2 G90677 cyn operon positiv
 265 17 77.3 299 2 C85528 cyn operon positiv
 266 17 77.3 300 2 H84022 arginase rocf [imp
 267 17 77.3 300 2 S15786 glucose transport
 268 17 77.3 304 2 B91113 hypothetical prote
 269 17 77.3 304 2 F70804 probable DNA GYCO
 270 17 77.3 304 2 B85958 hypothetical prote
 271 17 77.3 304 2 C65085 hypothetical prote
 272 17 77.3 308 2 S61448 heat shock transcr
 273 17 77.3 309 2 E87671 cytochrome c oxida
 274 17 77.3 310 2 A11059 carbamate kinase
 275 17 77.3 310 2 S49030 RNA-binding protei
 276 17 77.3 313 2 T21383 hypothetical prote
 277 17 77.3 313 2 C95247 conserved hypotet
 278 17 77.3 313 2 H98111 conserved hypotet
 279 17 77.3 314 2 T39510 probable mitochond
 280 17 77.3 314 2 S53492 RNA-binding protei
 281 17 77.3 316 2 AF0309 probable hydroxypr
 282 17 77.3 316 2 C88448 protein C45G9.5 [l
 283 17 77.3 317 1 B64123 probable dehydroge
 284 17 77.3 318 2 AD1709 mannose-6 phosphat
 285 17 77.3 319 2 C95927 probable sugar upt
 286 17 77.3 321 2 H81301 hypothetical prote
 287 17 77.3 322 1 JC4033 arginase (EC 3.5.3
 288 17 77.3 325 2 T24689 hypothetical prote
 289 17 77.3 328 2 AG2987 hypothetical prote
 290 17 77.3 329 2 H83099 probable oxidoredu
 291 17 77.3 329 2 T32760 hypothetical prote
 292 17 77.3 329 2 S28057 RNA-binding protei
 293 17 77.3 330 1 B64561 ketol-acid reducto
 294 17 77.3 330 2 A71945 ketol-acid reducto
 295 17 77.3 330 2 A83828 glucose kinase glk
 296 17 77.3 330 2 T37854 probable IUNH-fam
 297 17 77.3 331 2 AG0599 HlyD-family secret
 298 17 77.3 332 2 B85588 probable membrane
 299 17 77.3 332 2 A99738 probable membrane
 300 17 77.3 332 2 C64816 hypothetical prote

ALIGNMENTS

RESULT 1
 D82288
 NADPH-flavin oxidoreductase VC0715 [imported] - Vibrio cholerae (strain N16961 serogroup
 C;Species: Vibrio cholerae
 C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
 C;Accession: D82288
 R;Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
 chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, B
 1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
 Nature 406, 477-483, 2000
 A;Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
 A;Reference number: A82035; MUID:20406833; PMID:10952301
 A;Accession: D82288
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-240 <HEI>

A;Cross-references: UNIPROT:Q9KU15; GB:AE004158; GB:AE003852; NID:9655156; PIDN:AAF9388C
 A;Experimental source: serogroup O1; strain N16961; biotype El Tor
 C;Genetics:
 A;Gene: VC0715
 A;Map position: 1
 C;Superfamily: NADPH-flavin oxidoreductase homolog
 Query Match 86.4%; Score 19; DB 2; Length 240;
 Best Local Similarity 57.1%; Pred. No. 2.4e+02;
 Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 VXXXXAEF 7
 DB 70 VASAAEF 76
 RESULT 2
 F85885
 Probable regulator yfeU [imported] - Escherichia coli (strain O157:H7, substrain EDL933)
 C;Species: Escherichia coli
 C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
 C;Accession: F85885
 R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
 iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
 Nature 409, 529-533, 2001
 A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
 A;Reference number: A85480; MUID:21074935; PMID:11206551
 A;Accession: F85885
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-298 <STO>
 A;Cross-references: GB:AE005174; NID:gl2516806; PIDN:AAG57546.1; GSPDB:GN00145; UWGP:Z36
 A;Experimental source: strain O157:H7, substrain EDL933
 C;Genetics:
 A;Gene: yfeU
 C;Superfamily: conserved hypothetical protein b2428
 Query Match 86.4%; Score 19; DB 2; Length 298;
 Best Local Similarity 57.1%; Pred. No. 3e+02;
 Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 VXXXXAEF 7
 DB 172 VSTTAEF 178

RESULT 3
 C65017
 hypothetical protein b2428 - Escherichia coli (strain K-12)
 C;Species: Escherichia coli
 C;Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 09-Jul-2004
 C;Accession: C65017
 R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
 .A.; Rose, D.J.; Mau, B.; Shao, Y.
 Science 277, 1453-1462, 1997
 A;Title: The complete genome sequence of Escherichia coli K-12.
 A;Reference number: A64720; MUID:97426617; PMID:9278503
 A;Accession: C65017
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-298 <BLAT>
 A;Cross-references: UNIPROT:P76535; GB:AE000330; GB:U00096; NID:gl788763; PIDN:AAC75481.
 A;Experimental source: strain K-12, substrain MG1655
 C;Superfamily: conserved hypothetical protein b2428
 Query Match 86.4%; Score 19; DB 2; Length 298;
 Best Local Similarity 57.1%; Pred. No. 3e+02;
 Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 VXXXXAEF 7
 DB 172 VSTTAEF 178

```

RESULT 4
C:Probable regulator [imported] - Escherichia coli (strain O157:H7, substrain RIMD 0509952)
C:Species: Escherichia coli
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C:Accession: C91041
R:Hayaashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
Gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genetic
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: C91041
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-298 <HAY>
A:Cross-references: UNIPROT:Q8XBJ2; GB:BA000007; PIDN:BAR36722.1; PID:gl3362769; GSPDB:G
A:Experimental source: strain O157:H7, substrain RIMD 0509952
C:Genetics:
A:Gene: ECs3299
C:Superfamily: conserved hypothetical protein b2428

Query Match 86.4%; Score 19; DB 2; Length 298;
Best Local Similarity 57.1%; Pred. No. 3e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VXXAAEF 7
| | | |
Db 172 VSTTAEF 178

RESULT 5
A86165
protein F15K9.6 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C:Accession: A86165
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chen, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: A86165
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-385 <STO>
A:Cross-references: UNIPROT:Q9ZVT5; GB:AB005172; NID:g3850570; PIDN:AAC72110.1; GSPDB:GN
C:Genetics:
A:Gene: F15K9.6
A:Map Position: 1
C:Superfamily: Arabidopsis thaliana hypothetical protein T419.20

Query Match 86.4%; Score 19; DB 2; Length 385;
Best Local Similarity 57.1%; Pred. No. 3.9e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VXXAAEF 7
| | | |
Db 172 VSTAAEF 178

RESULT 6
T31224
traH protein homolog - Sphingomonas aromaticivorans plasmid pNL1
C:Species: Sphingomonas aromaticivorans
C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 09-Jul-2004
C:Accession: T31224

```

```

R;Romine, M.F.; Stillwell, L.C.; Wong, K.K.; Thurston, S.J.; Sisk, E.C.; Sensen, C.W.; Ga
submitted to the EMBL Data Library, July 1998
A:Description: Complete sequence of a 184 Kb catabolic plasmid from Sphingomonas aromatici
A:Reference number: Z20992
A:Accession: T31224
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-484 <ROM>
A:Cross-references: UNIPROT:O85932; EMBL:AF079317; NID:g3378261; PID:g3378365; PIDN:AAD0:
C:Genetics:
A:Genome: plasmid pNL1
A>Note: traH
C:Superfamily: traH protein

Query Match 86.4%; Score 19; DB 2; Length 484;
Best Local Similarity 57.1%; Pred. No. 4.9e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VXXAAEF 7
| | | |
Db 368 VSAAAEF 374

RESULT 7
AF0911
probable exported protein [imported] - Salmonella enterica subsp. enterica serovar Typhi
C:Species: Salmonella enterica subsp. enterica serovar Typhi
A:Note: this species has also been called Salmonella typhi
C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C:Accession: AF0911
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serovar
A:Reference number: AB0502; MUID:21534947; PMID:11677608
A:Accession: AF0911
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1266 <PAR>
A:Cross-references: GB:AL513382; PIDN:CAD07884.1; PID:g16504431; GSPDB:GN00176
C:Genetics:
A:Gene: STY3549
C:Superfamily: Escherichia coli hypothetical protein yHdP

Query Match 86.4%; Score 19; DB 2; Length 1266;
Best Local Similarity 57.1%; Pred. No. 1.3e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VXXAAEF 7
| | | |
Db 895 VSAAEF 901

RESULT 8
T00095
hypothetical protein KIAA0470 - human
C:Species: Homo sapiens (man)
C:Date: 22-Jan-1999 #sequence_revision 22-Jan-1999 #text_change 09-Jul-2004
C:Accession: T00095; T47155
R;Seki, N.; Ohira, M.; Nagase, T.; Ishikawa, K.; Miyajima, N.; Nakajima, D.; Nomura, N.;
DNA Res. 4, 345-349, 1997
A:Title: Characterization of cDNA clones in size-fractionated cDNA libraries from human
A:Reference number: Z14085; MUID:98116662; PMID:9455484
A:Accession: T00095
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1460 <SEK>
A:Cross-references: UNIPROT:O75058; EMBL:AB007939; NID:g3413901; PIDN:BAA32315.1; PID:g34
A:Experimental source: brain
R;Poustka, A.; Wellenreuther, R.; Mewes, H.W.; Weil, B.; Wiemann, S.
submitted to the Protein Sequence Database, March 2000

```

A;Reference number: 224374
A;Accession: T47155
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1288-1460 <AAA>
A;Cross-references: EMBL:ALJ161964
A;Experimental source: fetal brain; clone DKFZp564B0982
C;Genetics:
A;Map position: 1
A;Note: KIAA0470; DKFZp564B0982.1

Query Match 86.4%; Score 19; DB 2; Length 1460;
Best Local Similarity 57.1%; Pred. No. 1.5e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
| |||
Db 1425 VSAAAEF 1431

RESULT 9
AE1449
hypothetical protein lin0132 [imported] - Listeria innocua (strain Clip11262)
C;Species: Listeria innocua
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C;Accession: AE1449
R;Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A;Authors: Kraft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend,
A;Title: Comparative genomics of Listeria species.
A;Reference number: AB1077; MUID:21537279; PMID:11679669
A;Accession: AE1449
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1946 <GLA>
A;Cross-references: UNIPROT:Q92FH4; GB:AL592022; PIDN:CAC95365.1; PID:gl16412551; GSPDB:G
A;Experimental source: strain Clip11262
C;Genetics:
A;Gene: lin0132

Query Match 86.4%; Score 19; DB 2; Length 1946;
Best Local Similarity 57.1%; Pred. No. 2e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
| |||
Db 928 VSTAAEF 934

RESULT 10
T16305
hypothetical protein F40F4.6 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C;Accession: T16305
R;Wilson, R.

submitted to the EMBL Data Library, November 1995
A;Description: The sequence of C. elegans cosmid F40F4.

A;Reference number: Z18493

A;Accession: T16305

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-2214 <ML>

A;Cross-references: UNIPROT:Q20219; EMBL:U40420; NID:g1065513; PID:g1065514; PIDN:AAA814

C;Genetics:

A;Gene: CESP:F40F4.6

A;Introns: 57/3; 95/1; 302/3; 323/1; 380/1; 404/1; 468/1; 507/2; 547/3; 595/3; 1532/1; 1

Query Match 86.4%; Score 19; DB 2; Length 2214;
Best Local Similarity 57.1%; Pred. No. 2.2e+03;

Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 VXXAAEF 7
| |||
Db 1770 VTTSAEF 1776

RESULT 11

A54252

omega-agatoxin III, 8.478K - funnel-weaving spider (Agelenopsis aperta)

C;Species: Agelenopsis aperta

C;Date: 06-Oct-1994 #sequence_revision 18-Nov-1994 #text_change 19-Jan-1996

C;Accession: A54252

R;Ertel, E.A.; Warren, V.A.; Adams, M.E.; Griffin, P.R.; Cohen, C.J.; Smith, M.M.

Biochemistry 33, 5098-5108, 1994

A;Title: Type III omega-agatoxins: a family of probes for similar binding sites on L-an

A;Reference number: A54252; MUID:94227039; PMID:8172884

A;Accession: A54252

A;Status: preliminary

A;Molecule type: protein

A;Residues: 1-76 <ERT>

A;Experimental source: venom

A;Note: sequence extracted from NCBI backbone (NCBIP:148239)

Query Match 81.8%; Score 18; DB 2; Length 76;
Best Local Similarity 57.1%; Pred. No. 1.5e+02;

Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
| |||
Db 42 VGTSAEF 48

RESULT 12

D54252

omega-agatoxin III, 8.581K - funnel-weaving spider (Agelenopsis aperta)

C;Species: Agelenopsis aperta

C;Date: 06-Oct-1994 #sequence_revision 18-Nov-1994 #text_change 19-Jan-1996

C;Accession: D54252

R;Ertel, E.A.; Warren, V.A.; Adams, M.E.; Griffin, P.R.; Cohen, C.J.; Smith, M.M.

Biochemistry 33, 5098-5108, 1994

A;Title: Type III omega-agatoxins: a family of probes for similar binding sites on L-an

A;Reference number: A54252; MUID:94227039; PMID:8172884

A;Accession: D54252

A;Status: preliminary

A;Molecule type: protein

A;Residues: 1-76 <ERT>

A;Experimental source: venom

A;Note: sequence extracted from NCBI backbone (NCBIP:148242)

Query Match 81.8%; Score 18; DB 2; Length 76;
Best Local Similarity 57.1%; Pred. No. 1.5e+02;

Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
| |||
Db 42 VGTSAEF 48

RESULT 13

B54252

omega-agatoxin III, 8.637K - funnel-weaving spider (Agelenopsis aperta)

C;Species: Agelenopsis aperta

C;Date: 06-Oct-1994 #sequence_revision 18-Nov-1994 #text_change 19-Jan-1996

C;Accession: B54252

R;Ertel, E.A.; Warren, V.A.; Adams, M.E.; Griffin, P.R.; Cohen, C.J.; Smith, M.M.

Biochemistry 33, 5098-5108, 1994

A;Title: Type III omega-agatoxins: a family of probes for similar binding sites on L-an

A;Reference number: A54252; MUID:94227039; PMID:8172884

A;Accession: B54252

A;Status: preliminary

A;Molecule type: protein

A;Residues: 1-76 <ERT>

A;Experimental source: venom
A;Note: sequence extracted from NCBI backbone (NCBIP:148240)

Query Match 81.8%; Score 18; DB 2; Length 76;
Best Local Similarity 57.1%; Pred. No. 1.5e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXXAEEF 7
| |||
Db 42 VGTSAEEF 48

RESULT 14

C54252
omega-agatoxin III, 8.607K - funnel-weaving spider (Agelenopsis aperta)
C;Species: Agelenopsis aperta
C;Date: 06-Oct-1994 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C;Accession: C54252
R;Ertel, E.A.; Warren, V.A.; Adams, M.E.; Griffin, P.R.; Cohen, C.J.; Smith, M.M.
Biochemistry 33, 5098-5108, 1994
A;Title: Type III omega-agatoxins: a family of probes for similar binding sites on L- ar
A;Reference number: A54252; MUID:94227039; PMID:8172884
A;Accession: C54252
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-76 <ERT>
A;Cross-references: UNIPROT:P81744
A;Experimental source: venom
A;Note: sequence extracted from NCBI backbone (NCBIP:148241)

Query Match 81.8%; Score 18; DB 2; Length 76;
Best Local Similarity 57.1%; Pred. No. 1.5e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXXAEEF 7
| |||
Db 42 VGTSAEEF 48

RESULT 15

A42335
omega-agatoxin IIIA - funnel-weaving spider (Agelenopsis aperta)
C;Species: Agelenopsis aperta
C;Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C;Accession: A42335
R;Venema, V.J.; Swiderek, K.M.; Lee, T.D.; Hathaway, G.M.; Adams, M.E.
J. Biol. Chem. 267, 2610-2615, 1992
A;Title: Antagonism of synaptic calcium channels by subtypes of omega-agatoxins.
A;Reference number: A42335; MUID:92129351; PMID:1310319
A;Accession: A42335
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-76 <VEN>
A;Cross-references: UNIPROT:P33034
A;Experimental source: venom
A;Note: sequence extracted from NCBI backbone (NCBIP:78693)

Query Match 81.8%; Score 18; DB 2; Length 76;
Best Local Similarity 57.1%; Pred. No. 1.5e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXXAEEF 7
| |||
Db 42 VGTSAEEF 48

RESULT 16

A34520
calcium-binding protein, vitamin D-dependent - guinea pig (fragments)
C;Species: Cavia porcellus (guinea pig)
C;Date: 15-Jun-1990 #sequence_revision 15-Jun-1990 #text_change 09-Jul-2004
C;Accession: A34520
R;Winsky, L.; Nakata, H.; Martin, B.M.; Jacobowitz, D.M.

Proc. Natl. Acad. Sci. U.S.A. 86, 10139-10143, 1989
A;Title: Isolation, partial amino acid sequence, and immunohistochemical localization of
A;Reference number: A34520; MUID:90099311; PMID:2602362
A;Accession: A34520
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-92 <WIN>
A;Cross-references: UNIPROT:P20658
A;Superfamily: calretinin; calmodulin repeat homology
C;Keywords: brain; calcium binding; duplication; EF hand; vitamin D

Query Match 81.8%; Score 18; DB 2; Length 92;
Best Local Similarity 57.1%; Pred. No. 1.9e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXXAEEF 7
| |||
Db 30 VGSSAEEF 36

RESULT 17

S48420
probable membrane protein YIL059c - yeast (Saccharomyces cerevisiae)
C;Species: Saccharomyces cerevisiae
C;Date: 02-Dec-1994 #sequence_revision 02-Dec-1994 #text_change 09-Jul-2004
C;Accession: S48420
R;Smith, V.

submitted to the EMBL Data Library, September 1994

A;Reference number: S48407
A;Accession: S48420
A;Molecule type: DNA
A;Residues: 1-121 <SMI>
A;Cross-references: UNIPROT:P40520; GB:247047; EMBL:Z38060; NID:g603997; PID:g763287; GSI:
C;Genetics:
A;Gene: MIPS:YIL059c
A;Cross-references: SGD:S0001321
A;Map position: 9L
C;Superfamily: Saccharomyces probable membrane protein YIL059c
C;Keywords: transmembrane protein
F;6-22/Domain: transmembrane #status predicted <TMM>

Query Match 81.8%; Score 18; DB 2; Length 121;
Best Local Similarity 57.1%; Pred. No. 2.5e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXXAEEF 7
| |||
Db 19 VTALAEF 25

RESULT 18

T06471
core protein - garden pea
C;Species: Pisum sativum (garden pea)
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C;Accession: T06471
R;Pohlmeier, K.; Soll, J.; Steinkamp, T.; Wagner, R.
submitted to the EMBL Data Library, June 1996
A;Description: Isolation and characterisation of a new type of pore protein present in t
A;Reference number: Z15701
A;Accession: T06471
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-146 <POH>
A;Cross-references: UNIPROT:Q41050; EMBL:Z73553; PIDN:CAA97910.1
A;Experimental source: cv. Golf
C;Genetics:
A;Genome: nuclear
C;Keywords: chloroplast

Query Match 81.8%; Score 18; DB 2; Length 146;
Best Local Similarity 42.9%; Pred. No. 3e+02;
Matches 3; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
:
|||
Db 135 IATAAEF 141

RESULT 19

H84341
hypothetical protein Vng1907h [imported] - Halobacterium sp. NRC-1
C:Species: Halobacterium sp. NRC-1
C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C:Accession: H84341
R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Laaky, S.; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jabil
Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li
A:Title: Genome sequence of Halobacterium species NRC-1.
A:Reference number: A84160; MUID:20504483; PMID:11016950
A:Accession: H84341
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-147 <STO>
A:Cross-references: UNIPROT:Q9HNX0; GB:AE004437; NID:gl0581349; PIDN:AAG20100.1; GSPDB:G
C:Genetics:
A:Gene: VNG1907H

Query Match 81.8%; Score 18; DB 2; Length 147;
Best Local Similarity 57.1%; Pred. No. 3e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
:
|||
Db 102 VSSRAEF 108

RESULT 20

A86878
non-heme iron-binding ferritin [imported] - Lactococcus lactis subsp. lactis (strain IL1
C:Species: Lactococcus lactis subsp. lactis
C>Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 09-Jul-2004
C:Accession: A86878
R;Bolotin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarne, K.; Weissbach, J.; Ehrli
Genome Res. 11, 731-753, 2001
A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis se
A:Reference number: A86625; MUID:21235186; PMID:11337471
A:Accession: A86878
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-148 <STO>
A:Cross-references: UNIPROT:Q9CE23; GB:AE005176; PID:gl2725073; PIDN:AAK06123.1; GSPDB:G
A:Experimental source: strain IL1403
C:Genetics:
A:Gene: dpsA
C:Superfamily: hypothetical protein H11349

Query Match 81.8%; Score 18; DB 2; Length 148;
Best Local Similarity 57.1%; Pred. No. 3e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
:
|||
Db 69 VSTLAEF 75

RESULT 21

C87486
hypothetical protein CC1912 [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C>Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C:Accession: C87486
R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon

n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: A87249; MUID:21173698; PMID:11259647
A:Accession: C87486
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-159 <STO>
A:Cross-references: UNIPROT:Q9A714; GB:AE005673; NID:gl3423365; PIDN:AAK23887.1; GSPDB:G
C:Genetics:
A:Gene: CC1912
C:Superfamily: (3R)-hydroxymyristoyl- [acyl carrier protein] dehydratase

Query Match 81.8%; Score 18; DB 2; Length 159;
Best Local Similarity 57.1%; Pred. No. 3.3e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
:
|||
Db 142 VAAAEF 148

RESULT 22

S41476
calretinin - mouse (fragment)
C:Species: Mus musculus (house mouse)
C>Date: 13-Jan-1995 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C:Accession: S41476; S34391
R;Ellis, J.; Rogers, J.
Nucleic Acids Res. 21, 5171-5178, 1993
A:Title: Design and specificity of hammerhead ribozymes against calretinin mRNA.
A:Reference number: S41476; MUID:9407721; PMID:8255773
A:Accession: S41476
A:Molecule type: mRNA
A:Residues: 1-190 <ELL>
A:Cross-references: UNIPROT:Q08331; EMBL:X73985; NID:g939386; PIDN:CAA52163.1; PID:g939338
C:Superfamily: calretinin; calmodulin repeat homology
C:Keywords: brain; calcium binding; duplication; EF hand
F;26-58/Domain: calmodulin repeat homology <EF3>
F;70-102/Domain: calmodulin repeat homology <EF4>
F;114-146/Domain: calmodulin repeat homology <EF5>

Query Match 81.8%; Score 18; DB 2; Length 190;
Best Local Similarity 57.1%; Pred. No. 3.9e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
:
|||
Db 25 VGSSAEF 31

RESULT 23

T34819
hypothetical protein SC2E9.08 - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C>Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 09-Jul-2004
C:Accession: T34819
R;Oliver, K.; Harris, D.; Parkhill, J.; Barrall, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, January 1998
A:Reference number: Z21558
A:Accession: T34819
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-194 <OLI>
A:Cross-references: UNIPROT:O54133; EMBL:AL021530; PIDN:CAA16476.1; GSPDB:GN000070; SCOE
A:Experimental source: strain A3(2)
C:Genetics:
A:Gene: SCOE9.08
C:Superfamily: Streptomyces coelicolor hypothetical protein SC2E9.08

Query Match 81.8%; Score 18; DB 2; Length 194;
Best Local Similarity 57.1%; Pred. No. 4e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
|
|
|
Db 41 VEAARF 47

RESULT 29

S65426
pyruvate decarboxylase (EC 4.1.1.1) - fava bean (fragment)
C;Species: Vicia faba (fava bean)
C;Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 12-Jul-2004
C;Accession: S65426
R;Muecke, U.; Wohlfiarth, T.; Fiedler, U.; Baumelein, H.; Ruecknagel, K.P.; Koenig, S.
Eur. J. Biochem. 237, 373-382, 1996
A;Title: Pyruvate decarboxylase from Pisum sativum. Properties, nucleotide and amino acid
A;Reference number: S65423; MUID:96215432; PMID:8647075
A;Accession: S65426
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-228 <MUE>
A;Cross-references: UNIPROT:Q7M228
C;Superfamily: thiamin pyrophosphate-binding domain homology
C;Keywords: carbon-carbon lyase; carboxy-lyase

Query Match 81.8%; Score 18; DB 2; Length 228;
Best Local Similarity 57.1%; Pred. No. 4.7e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
|
|
|
Db 41 VEAARF 47

RESULT 30

AG0605
oxygen-insensitive NADPH nitroreductase (EC 1.-.-.-) [imported] - Salmonella enterica su
C;Species: Salmonella enterica subsp. enterica serovar Typhi
A;Note: this species has also been called Salmonella typhi
C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C;Accession: AG0605
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A;Reference number: AB0502; MUID:21534947; PMID:11677608
A;Accession: AG0605
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-240 <PAR>
A;Cross-references: GB:AL513382; PIDN:CAD05313.1; PID:gl6502077; GSPDB:GN00176
C;Genetics:
A;Gene: nfaA
C;Superfamily: NADPH-flavin oxidoreductase homolog
C;Keywords: oxidoreductase

Query Match 81.8%; Score 18; DB 2; Length 240;
Best Local Similarity 57.1%; Pred. No. 4.9e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
|
|
|
Db 70 VAQAAEF 76

RESULT 31

180318
drug activity modulator A - Escherichia coli (strain K-12)
N;Contains: aromatic nitrate reductase (NADPH) (EC 1.6.6.-) oxygen-insensitive
C;Species: Escherichia coli
C;Date: 07-Jun-1996 #sequence_revision 07-Jun-1996 #text_change 09-Jul-2004
C;Accession: 180318; C64823; S04774

R;Chatterjee, P.K.; Sternberg, N.L.
Proc. Natl. Acad. Sci. U.S.A. 92, 8950-8954, 1995
A;Title: A general genetic approach in Escherichia coli for determining the mechanism(s)
A;Reference number: 159418; MUID:96004656; PMID:7568050
A;Accession: 180318
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-240 <RES>
A;Cross-references: UNIPROT:P17117; EMBL:U18655; NID:9609323; PIDN:AAC43450.1; PID:96093;
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Choi,
A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A;Title: The complete genome sequence of Escherichia coli K-12.
A;Reference number: A64720; MUID:97426617; PMID:9278503
A;Accession: C64823
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-240 <BLAT>
A;Cross-references: GB:AE000187; GB:U00096; NID:91787070; PIDN:AAC73938.1; PID:91787075;
A;Experimental source: strain K-12, substrain MG1655
R;Kang, W.K.; Icho, T.; Isono, S.; Kitakawa, M.; Isono, K.
Mol. Gen. Genet. 217, 281-289, 1989
A;Title: Characterization of the gene rimK responsible for the addition of glutamic acid
A;Reference number: S04774; MUID:89364710; PMID:2570347
A;Accession: S04774
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 134-154, 'CA', 157, 'AGH', 161-186, 'RHWRSMTSNWNITSPVAIIAGIPGAIISAEQSLKKAHLFWI
C;Genetics:
A;Gene: mdxA; mdx18
A;Map position: 18.7-19.0 min
C;Superfamily: NADPH-flavin oxidoreductase homolog
C;Keywords: flavoprotein; FMN; oxidoreductase

Query Match 81.8%; Score 18; DB 2; Length 240;
Best Local Similarity 57.1%; Pred. No. 4.9e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
|
|
|
Db 70 VAQAAEF 76

RESULT 32

C90745
modulator of drug activity A [imported] - Escherichia coli (strain O157:H7, substrain RIN
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C;Accession: C90745
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.;
gaawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genom
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: C90745
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-240 <HAY>
A;Cross-references: UNIPROT:Q8X6S1; GB:BA000007; PIDN:BA34354.1; PID:gl13360390; GSPDB:G
A;Experimental source: strain O157:H7, substrain RIMD 0509952
C;Genetics:
A;Gene: EC0931
C;Superfamily: NADPH-flavin oxidoreductase homolog

Query Match 81.8%; Score 18; DB 2; Length 240;
Best Local Similarity 57.1%; Pred. No. 4.9e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
|
|
|
Db 70 VAQAAEF 76

```

RESULT 33
G85595
modulator of drug activity A [imported] - Escherichia coli (strain O157:H7, substrain ED
C/Species: Escherichia coli
C/Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C/Accession: G85595
R/Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamouisis, K.; Apodaca,
Nature 409, 529-533, 2001
A/Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A/Reference number: A85480; MUID:21074935; PMID:11206551
A/Accession: G85595
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-240 <NAK>
A/Cross-references: UNIPROT:Q8X6S1; GB:AE005174; NID:q12513866; PIDN:AAG55227.1; GSPDB:G
A/Experimental source: strain O157:H7, substrain EDL933
C/Genetics:
A/Gene: mdxA
C/Superfamily: NADPH-flavin oxidoreductase homolog

Query Match      81.8%; Score 18; DB 2; Length 240;
Best Local Similarity 57.1%; Pred. No. 4.9e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VXXAAEF 7
| | | |
Db 70 VAQAAEF 76

RESULT 34
PQ0662
outer capsid spike protein VP4 - feline rotavirus A (strain Cat2) (fragment)
N/Alternate names: VP8* protein
C/Species: feline rotavirus A
C/Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 20-Jun-2000
C/Accession: PQ0662
R/Nakagomi, O.; Isegawa, Y.; Ueda, S.; Gerna, G.; Sarasini, A.; Kaga, E.; Nakagomi, T.;
J. Gen. Virol. 74, 1709-1713, 1993
A/Title: Nucleotide sequence comparison of the VP8* gene of rotaviruses possessing the A
A/Reference number: PQ0651; MUID:93346985; PMID:8393919
A/Accession: PQ0662
A/Molecule type: mRNA
A/Residues: 1-247 <NAK>
A/Cross-references: GB:D14621; NID:g287424; PIDN:BAA03472.1; PID:g287425
C/Comment: This protein plays a role in viral neutralization, viral virulence and haemag
C/Genetics:
A/Gene: VP4
C/Superfamily: rotavirus outer layer protein VP3

Query Match      81.8%; Score 18; DB 2; Length 247;
Best Local Similarity 57.1%; Pred. No. 5.1e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VXXAAEF 7
| | | |
Db 199 VSSDAEF 205

RESULT 35
PQ0652
outer capsid spike protein VP4 - human rotavirus (strain AU228) (fragment)
N/Alternate names: VP8* protein
C/Species: human rotavirus
C/Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 20-Jun-2000
C/Accession: PQ0652
R/Nakagomi, O.; Isegawa, Y.; Ueda, S.; Gerna, G.; Sarasini, A.; Kaga, E.; Nakagomi, T.;
J. Gen. Virol. 74, 1709-1713, 1993
A/Title: Nucleotide sequence comparison of the VP8* gene of rotaviruses possessing the A
A/Reference number: PQ0651; MUID:93346985; PMID:8393919
A/Accession: PQ0652
A/Molecule type: mRNA

```

```

A/Residues: 1-247 <NAK>
A/Cross-references: GB:D14615; NID:g287412; PIDN:BAA03466.1; PID:g287413
C/Comment: This protein plays a role in viral neutralization, viral virulence and haemag
C/Genetics:
A/Gene: VP4
C/Superfamily: rotavirus outer layer protein VP3
C/Keywords: hemagglutinin; outer capsid protein

Query Match      81.8%; Score 18; DB 2; Length 247;
Best Local Similarity 57.1%; Pred. No. 5.1e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VXXAAEF 7
| | | |
Db 199 VSSDAEF 205

RESULT 36
PQ0661
outer capsid spike protein VP4 - human rotavirus (strain MZ58) (fragment)
C/Species: human rotavirus
C/Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 20-Jun-2000
C/Accession: PQ0661
R/Nakagomi, O.; Isegawa, Y.; Ueda, S.; Gerna, G.; Sarasini, A.; Kaga, E.; Nakagomi, T.;
J. Gen. Virol. 74, 1709-1713, 1993
A/Title: Nucleotide sequence comparison of the VP8* gene of rotaviruses possessing the A
A/Reference number: PQ0651; MUID:93346985; PMID:8393919
A/Accession: PQ0661
A/Molecule type: mRNA
A/Residues: 1-247 <NAK>
A/Cross-references: GB:D14622; NID:g287426; PIDN:BAA03473.1; PID:g287427
C/Comment: This protein plays a role in viral neutralization, viral virulence and haemag
C/Genetics:
A/Gene: VP4
C/Superfamily: rotavirus outer layer protein VP3
C/Keywords: hemagglutinin; outer capsid protein

Query Match      81.8%; Score 18; DB 2; Length 247;
Best Local Similarity 57.1%; Pred. No. 5.1e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VXXAAEF 7
| | | |
Db 199 VSSDAEF 205

RESULT 37
PQ0655
outer capsid spike protein VP4 - human rotavirus (strain AU379) (fragment)
N/Alternate names: VP8* protein
C/Species: human rotavirus
C/Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 20-Jun-2000
C/Accession: PQ0655
R/Nakagomi, O.; Isegawa, Y.; Ueda, S.; Gerna, G.; Sarasini, A.; Kaga, E.; Nakagomi, T.;
J. Gen. Virol. 74, 1709-1713, 1993
A/Title: Nucleotide sequence comparison of the VP8* gene of rotaviruses possessing the A
A/Reference number: PQ0651; MUID:93346985; PMID:8393919
A/Accession: PQ0655
A/Molecule type: mRNA
A/Residues: 1-247 <NAK>
A/Cross-references: GB:D14616; NID:g287414; PIDN:BAA03467.1; PID:g287415
C/Comment: This protein plays a role in viral neutralization, viral virulence and haemag
C/Genetics:
A/Gene: VP4
C/Superfamily: rotavirus outer layer protein VP3
C/Keywords: hemagglutinin; outer capsid protein

Query Match      81.8%; Score 18; DB 2; Length 247;
Best Local Similarity 57.1%; Pred. No. 5.1e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VXXAAEF 7
| | | |
Db 199 VSSDAEF 205

```



```
Db      199 VSSDAEF 205

RESULT 38
PQ0651
outer capsid spike protein VP4 - human rotavirus (strain AU1115) (fragment)
N;Alternate names: VP8* protein
C;Species: human rotavirus
C;Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 20-Jun-2000
C;Accession: PQ0651
R;Nakagomi, O.; Isegawa, Y.; Ueda, S.; Gerna, G.; Sarasini, A.; Kaga, E.; Nakagomi, T.;
J. Gen. Virol. 74, 1709-1713, 1993
A;Title: Nucleotide sequence comparison of the VP8* gene of rotaviruses possessing the A
A;Reference number: PQ0651; MUID:93346985; PIDN:8393919
A;Accession: PQ0651
A;Molecule type: mRNA
A;Residues: 1-247 <NAK>
A;Cross-references: GB:D14613; NID:G287408; PIDN:BAA03464.1; PID:G287409
A;Note: the authors translated the codon ATG for residue 106 as Val and GTT for residue
C;Comment: This protein plays a role in viral neutralization, viral virulence and haemag
C;Genetics:
A;Gene: VP4
C;Superfamily: rotavirus outer layer protein VP3
C;Keywords: hemagglutinin; outer capsid protein

Query Match      81.8%; Score 18; DB 2; Length 247;
Best Local Similarity 57.1%; Pred. No. 5.1e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 VXXAAEF 7
      | |||
Db      199 VSSDAEF 205

RESULT 39
PQ0656
outer capsid spike protein VP4 - human rotavirus (strain AU387) (fragment)
N;Alternate names: VP8* protein
C;Species: human rotavirus
C;Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 20-Jun-2000
C;Accession: PQ0656
R;Nakagomi, O.; Isegawa, Y.; Ueda, S.; Gerna, G.; Sarasini, A.; Kaga, E.; Nakagomi, T.;
J. Gen. Virol. 74, 1709-1713, 1993
A;Title: Nucleotide sequence comparison of the VP8* gene of rotaviruses possessing the A
A;Reference number: PQ0651; MUID:93346985; PIDN:8393919
A;Accession: PQ0656
A;Molecule type: mRNA
A;Residues: 1-247 <NAK>
A;Cross-references: GB:D14617; NID:G287416; PIDN:BAA03468.1; PID:G287417
A;Comment: This protein plays a role in viral neutralization, viral virulence and haemag
C;Genetics:
A;Gene: VP4
C;Superfamily: rotavirus outer layer protein VP3
C;Keywords: hemagglutinin; outer capsid protein

Query Match      81.8%; Score 18; DB 2; Length 247;
Best Local Similarity 57.1%; Pred. No. 5.1e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 VXXAAEF 7
      | |||
Db      199 VSSDAEF 205

RESULT 40
PQ0660
outer capsid spike protein VP4 - human rotavirus (strain PCP5) (fragment)
N;Alternate names: VP8* protein
C;Species: human rotavirus
C;Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 20-Jun-2000
C;Accession: PQ0660
R;Nakagomi, O.; Isegawa, Y.; Ueda, S.; Gerna, G.; Sarasini, A.; Kaga, E.; Nakagomi, T.;
J. Gen. Virol. 74, 1709-1713, 1993
A;Title: Nucleotide sequence comparison of the VP8* gene of rotaviruses possessing the A
A;Reference number: PQ0651; MUID:93346985; PIDN:8393919
A;Accession: PQ0651
A;Molecule type: mRNA
A;Residues: 1-247 <NAK>
A;Cross-references: GB:D14624; NID:G287430; PIDN:BAA03475.1; PID:G287431
A;Comment: This protein plays a role in viral neutralization, viral virulence and haemag
C;Genetics:
A;Gene: VP4
C;Superfamily: rotavirus outer layer protein VP3
C;Keywords: hemagglutinin; outer capsid protein

Query Match      81.8%; Score 18; DB 2; Length 247;
Best Local Similarity 57.1%; Pred. No. 5.1e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 VXXAAEF 7
      | |||
Db      199 VSSDAEF 205

RESULT 41
PQ0659
outer capsid spike protein VP4 - human rotavirus (strain PA151) (fragment)
N;Alternate names: VP8* protein
C;Species: human rotavirus
C;Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 20-Jun-2000
C;Accession: PQ0659
R;Nakagomi, O.; Isegawa, Y.; Ueda, S.; Gerna, G.; Sarasini, A.; Kaga, E.; Nakagomi, T.;
J. Gen. Virol. 74, 1709-1713, 1993
A;Title: Nucleotide sequence comparison of the VP8* gene of rotaviruses possessing the A
A;Reference number: PQ0651; MUID:93346985; PIDN:8393919
A;Accession: PQ0659
A;Molecule type: mRNA
A;Residues: 1-247 <NAK>
A;Cross-references: GB:D14623; NID:G287428; PIDN:BAA03474.1; PID:G287429
A;Comment: This protein plays a role in viral neutralization, viral virulence and haemag
C;Genetics:
A;Gene: VP4
C;Superfamily: rotavirus outer layer protein VP3
C;Keywords: hemagglutinin; outer capsid protein

Query Match      81.8%; Score 18; DB 2; Length 247;
Best Local Similarity 57.1%; Pred. No. 5.1e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 VXXAAEF 7
      | |||
Db      199 VSSDAEF 205

RESULT 42
PQ0658
outer capsid spike protein VP4 - human rotavirus (strain AU938) (fragment)
N;Alternate names: VP8* protein
C;Species: human rotavirus
C;Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 20-Jun-2000
C;Accession: PQ0658
R;Nakagomi, O.; Isegawa, Y.; Ueda, S.; Gerna, G.; Sarasini, A.; Kaga, E.; Nakagomi, T.;
J. Gen. Virol. 74, 1709-1713, 1993
A;Title: Nucleotide sequence comparison of the VP8* gene of rotaviruses possessing the A
A;Reference number: PQ0651; MUID:93346985; PIDN:8393919
A;Accession: PQ0658
A;Molecule type: mRNA
A;Residues: 1-247 <NAK>
A;Cross-references: GB:D14620; NID:G287422; PIDN:BAA03471.1; PID:G287423
A;Comment: This protein plays a role in viral neutralization, viral virulence and haemag
C;Genetics:
A;Gene: VP4
C;Superfamily: rotavirus outer layer protein VP3
C;Keywords: hemagglutinin; outer capsid protein

Query Match      81.8%; Score 18; DB 2; Length 247;
```

Best Local Similarity 57.1%; Pred. No. 5.1e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VXXXXAEF 7
|
199 VSSDAEF 205

RESULT 43

PQ0653
outer capsid spike protein VP4 - human rotavirus (strain AU125) (fragment)
N;Alternate names: VP8* protein
C;Species: human rotavirus
C;Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 20-Jun-2000
R;Accession: PQ0653
R;Nakagomi, O.; Isegawa, Y.; Ueda, S.; Gerna, G.; Sarasini, A.; Kaga, E.; Nakagomi, T.; J. Gen. Virol. 74, 1709-1713, 1993
A;Title: Nucleotide sequence comparison of the VP8* gene of rotaviruses possessing the A
A;Reference number: PQ0651; MUID:93346985; PMID:8393919
A;Accession: PQ0653
A;Molecule type: mRNA
A;Residues: 1-247 <NAK>
A;Cross-references: GB:D14614; NID:G287410; PIDN:BAA03465.1; PID:G287411
C;Comment: This protein plays a role in viral neutralization, viral virulence and haemagglutination
C;Genetics:
A;Gene: VP4
C;Superfamily: rotavirus outer layer protein VP3
C;Keywords: hemagglutinin; outer capsid protein

Query Match 81.8%; Score 18; DB 2; Length 247;
Best Local Similarity 57.1%; Pred. No. 5.1e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VXXXXAEF 7
|
199 VSSDAEF 205

RESULT 44

PQ0657
outer capsid spike protein VP4 - human rotavirus A (strain AU785) (fragment)
N;Alternate names: VP8* protein
C;Species: human rotavirus A
C;Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 20-Jun-2000
R;Accession: PQ0657; PQ0654
R;Nakagomi, O.; Isegawa, Y.; Ueda, S.; Gerna, G.; Sarasini, A.; Kaga, E.; Nakagomi, T.; J. Gen. Virol. 74, 1709-1713, 1993
A;Title: Nucleotide sequence comparison of the VP8* gene of rotaviruses possessing the A
A;Reference number: PQ0651; MUID:93346985; PMID:8393919
A;Accession: PQ0657
A;Molecule type: mRNA
A;Residues: 1-247 <NAK1>
A;Cross-references: GB:D14619; NID:G287420; PIDN:BAA03470.1; PID:G287421
A;Experimental source: strain AU785
A;Accession: PQ0654
A;Molecule type: mRNA
A;Residues: 1-247 <NAK2>
A;Cross-references: GB:D14618; NID:G287418; PIDN:BAA03469.1; PID:G287419
A;Experimental source: strain AU720
C;Comment: This protein plays a role in viral neutralization, viral virulence and haemagglutination
C;Genetics:
A;Gene: VP4
C;Superfamily: rotavirus outer layer protein VP3
C;Keywords: hemagglutinin; outer capsid protein

Query Match 81.8%; Score 18; DB 2; Length 247;
Best Local Similarity 57.1%; Pred. No. 5.1e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VXXXXAEF 7
|
199 VSSDAEF 205

RESULT 45

S61221
outer capsid spike protein VP-8 - human rotavirus A (isolate 5193) (fragment)
N;Alternate names: VP-8 protein
C;Species: human rotavirus A
A;Variety: isolate 5193
C;Date: 18-Sep-1997 #sequence_revision 18-Sep-1997 #text_change 09-Jul-2004
C;Accession: S61221
R;Gollop, R.; Kaga, E.; Silberstein, I.; Shulman, L.; Nakagomi, O.; Mendelson, E.; Shif, submitted to the EMBL Data Library, August 1995
A;Description: Three forms of AU-1 like viruses differentiated by their overall genomic
A;Reference number: S61221
A;Accession: S61221
A;Molecule type: Genomic RNA
A;Residues: 1-247 <GOL>
A;Cross-references: UNIPROT:Q86509; EMBL:X90733; NID:G967070; PIDN:CAA62267.1; PID:G967070;
A;Experimental source: strain AU-1; isolate 5193
C;Genetics:
A;Gene: VP-8
C;Superfamily: rotavirus outer layer protein VP3
C;Keywords: outer capsid protein

Query Match 81.8%; Score 18; DB 2; Length 247;
Best Local Similarity 57.1%; Pred. No. 5.1e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VXXXXAEF 7
|
199 VSSDAEF 205

RESULT 46

S61224
outer capsid spike protein VP-8 - human rotavirus A (isolate 6460 and others) (fragment)
N;Alternate names: VP-8 protein
C;Species: human rotavirus A
A;Variety: isolate 6460; isolate 6784; isolate 6584; isolate 7044
C;Date: 18-Sep-1997 #sequence_revision 18-Sep-1997 #text_change 09-Jul-2004
C;Accession: S61224; S61225; S61226; S61227
R;Gollop, R.; Kaga, E.; Silberstein, I.; Shulman, L.; Nakagomi, O.; Mendelson, E.; Shif, submitted to the EMBL Data Library, August 1995
A;Description: Three forms of AU-1 like viruses differentiated by their overall genomic
A;Reference number: S61221
A;Accession: S61224
A;Molecule type: Genomic RNA
A;Residues: 1-247 <GOL>
A;Cross-references: UNIPROT:Q86512; EMBL:X90736; NID:G967076; PIDN:CAA62270.1; PID:G967070;
A;Experimental source: strain AU-1; isolate 6460
A;Accession: S61225
A;Molecule type: Genomic RNA
A;Residues: 1-247 <GOW>
A;Cross-references: EMBL:X90738
A;Experimental source: strain AU-1; isolate 6784
A;Accession: S61226
A;Molecule type: Genomic RNA
A;Residues: 1-247 <GOF>
A;Cross-references: EMBL:X90737; NID:G967078; PIDN:CAA62271.1; PID:G967079
A;Experimental source: strain AU-1; isolate 6584
A;Accession: S61227
A;Molecule type: Genomic RNA
A;Residues: 1-247 <GOA>
A;Cross-references: EMBL:X90739; NID:G967082; PIDN:CAA62273.1; PID:G967083
A;Experimental source: strain AU-1; isolate 7044
C;Genetics:
A;Gene: VP-8
C;Superfamily: rotavirus outer layer protein VP3
C;Keywords: outer capsid protein

Query Match 81.8%; Score 18; DB 2; Length 247;
Best Local Similarity 57.1%; Pred. No. 5.1e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VXXAAEF 7
|
199 VSSDAEF 205

RESULT 47

S61222
outer capsid spike protein VP-8 - human rotavirus A (isolate 5829 and isolate 5960) (fragment)
N;Alternate names: VP-8 protein
C;Species: human rotavirus A
A;Variety: isolate 5829; isolate 5960
C;Date: 18-Sep-1997 #sequence_revision 18-Sep-1997 #text_change 09-Jul-2004
C;Accession: S61222; S61223
R;Gollop, R.; Kaga, E.; Silberstein, I.; Shulman, L.; Nakagomi, O.; Mendelson, E.; Shif.,
submitted to the EMBL Data Library, August 1995
A;Description: Three forms of AU-1 like viruses differentiated by their overall genomic
A;Reference number: S61221
A;Accession: S61222
A;Molecule type: genomic RNA
A;Residues: 1-247 <GOL>
A;Cross-references: UNIPROT:Q86510; EMBL:X90734; NID:g967072; PIDN:CAA62268.1; PID:g967070
A;Experimental source: strain AU-1; isolate 5829
A;Accession: S61223
A;Molecule type: genomic RNA
A;Residues: 1-247 <GOW>
A;Cross-references: EMBL:X90735; NID:g967074; PIDN:CAA62269.1; PID:g967075
A;Experimental source: strain AU-1; isolate 5960
C;Genetics:
A;Gene: VP-8
C;Superfamily: rotavirus outer layer protein VP3
C;Keywords: outer capsid protein

Query Match 81.8%; Score 18; DB 2; Length 247;
Best Local Similarity 57.1%; Pred. No. 5.1e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VXXAAEF 7
|
199 VSSDAEF 205

RESULT 48

S74373
heat shock protein grpE - Synecchocystis sp. (strain PCC 6803)
N;Alternate names: hypothetical protein sl10057
C;Species: Synecchocystis sp.
A;Variety: PCC 6803
C;Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
C;Accession: S74373
R;Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
DNA Res. 3, 109-136, 1996
A;Title: Sequence analysis of the genome of the unicellular cyanobacterium Synecchocystis
s.
A;Reference number: S74322; MUID:97061201; PMID:8905231
A;Accession: S74373
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-249 <KAN>
A;Cross-references: UNIPROT:Q59978; EMBL:D64001; GB:AB001339; NID:g1001102; PIDN:BAAL029
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C;Genetics:
A;Gene: grpE
C;Superfamily: heat shock protein grpE

Query Match 81.8%; Score 18; DB 2; Length 249;
Best Local Similarity 57.1%; Pred. No. 5.1e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VXXAAEF 7
|
101 VALAAEF 107

RESULT 49

S34724
Probable oxidoreductase (EC 1.1.1.1.-) - Serratia marcescens
C;Species: Serratia marcescens
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004
C;Accession: S34724
R;Viejo, M.; Enfedaque, J.; Regue, M.
submitted to the EMBL Data Library, July 1993
A;Description: Bacteriocin 28b from Serratia marcescens does not present lysis nor immun
A;Reference number: S34724
A;Accession: S34724
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-254 <VIE>
A;Cross-references: UNIPROT:Q54472; EMBL:Z25281; NID:g395948; PIDN:CAA80898.1; PID:g39594
C;Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology
C;Keywords: oxidoreductase
F;7-190/Domain: short-chain alcohol dehydrogenase homology <SADH>

Query Match 81.8%; Score 18; DB 2; Length 254;
Best Local Similarity 57.1%; Pred. No. 5.2e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VXXAAEF 7
|
75 VSAGAEF 81

RESULT 50

I40886
glycine hydroxymethyltransferase (EC 2.1.2.1) - Corynebacterium sp. (fragment)
N;Alternate names: serine aldolase; serine hydroxymethylase; threonine aldolase
C;Species: Corynebacterium sp.
C;Date: 18-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 27-Oct-2003
C;Accession: I40886
R;Chlumsky, L.J.; Zhang, L.; Jorns, M.S.
J. Biol. Chem. 270, 18252-18259, 1995
A;Title: Sequence analysis of sarcosine oxidase and nearby genes reveals homologies with
A;Reference number: A57385; MUID:95355441; PMID:7543100
A;Accession: I40886
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-260 <RES>
A;Cross-references: EMBL:U23955; NID:g927587; PIDN:AAC43458.1; PID:g927588
C;Genetics:
A;Gene: glyA

C;Superfamily: serine/glycine hydroxymethyltransferase
C;Keywords: phosphoprotein; pyridoxal phosphate; transferase
F;60/Binding site: pyridoxal phosphate (Lys) (covalent) #status predicted

Query Match 81.8%; Score 18; DB 2; Length 260;
Best Local Similarity 42.9%; Pred. No. 5.3e+02;
Matches 3; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VXXAAEF 7
|
107 IAASAEF 113

RESULT 51

F69337
thiamin biosynthesis protein thil homolog - Archaeoglobus fulgidus
C;Species: Archaeoglobus fulgidus
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C;Accession: F69337
R;Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson
; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, B.F.
Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
Nature 390, 364-370, 1997
A;Authors: Uitterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.;
Smith, H.O.; Woese, C.R.; Venter, J.C.
A;Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo

A;Reference number: A69250; MUID:98049343; PMID:9389475

A;Accession: F69337
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-260 <KLE>
A;Cross-references: UNIPROT:O29556; GB:AE001056; GB:AE000782; NID:g2689379; PIDN:AAB9053
C;Superfamily: thiamin biosynthesis protein thl

Query Match 81.8%; Score 18; DB 2; Length 260;
Best Local Similarity 57.1%; Pred. No. 5.3e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
| |||
Db 104 VADSAEF 110

RESULT 52

S50699
orotidine-5'-phosphate decarboxylase (EC 4.1.1.23) - yeast (Pichia ohmeri)
C;Species: Pichia ohmeri
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 09-Jul-2004
C;Accession: S50699
R;Piredda, S.; Gaillardin, C.
Yeast 10, 1601-1612, 1994
A;Title: Development of a transformation system for the yeast *Yamadazyma* (Pichia) ohmeri
A;Reference number: S50698; MUID:95242835; PMID:7725795
A;Accession: S50699
A;Molecule type: DNA
A;Residues: 1-261 <PIR>
A;Cross-references: UNIPROT:P48844; EMBL:Z35100; NID:g510926; PIDN:CAA84483.1; PID:g5109
A;Note: the source was formerly designated as Pichia ohmeri
C;Genetics:
A;Gene: URA3
C;Superfamily: orotidine-5'-phosphate decarboxylase; orotidine-5'-phosphate decarboxylase
C;Keywords: carbon-carbon lyase; carboxy-lyase; pyrimidine nucleotide biosynthesis
F;1-261/Domain: orotidine-5'-phosphate decarboxylase homology <OPD>

Query Match 81.8%; Score 18; DB 2; Length 261;
Best Local Similarity 57.1%; Pred. No. 5.4e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
| |||
Db 35 VKTAEF 41

RESULT 53

S19184
hypothetical protein - Neisseria gonorrhoeae
C;Species: Neisseria gonorrhoeae
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004
C;Accession: S19184
R;Lavitola, A.; Vanni, M.; Martin, M.V.; Bruni, C.B.
submitted to the EMBL Data Library, January 1992
A;Description: Cloning and characterization of a *Neisseria* gene homologous to hisJ and a
A;Reference number: S19184
A;Accession: S19184
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-268 <LAV>
A;Cross-references: UNIPROT:Q06758; EMBL:X64421; NID:g49067; PIDN:CAA45768.1; PID:g49068
C;Superfamily: lysine-arginine-ornithine-binding protein

Query Match 81.8%; Score 18; DB 2; Length 268;
Best Local Similarity 57.1%; Pred. No. 5.5e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
| |||
Db 41 VASNAEF 47

RESULT 54

H81806
histidine-binding periplasmic protein NMA1811 [imported] - *Neisseria meningitidis* (strain
C;Species: *Neisseria meningitidis*
C;Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 09-Jul-2004
C;Accession: H81806
R;Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel
; Holroyd, S.; Jagers, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,
Nature 404, 502-506, 2000
A;Title: Complete DNA sequence of a serogroup A strain of *Neisseria meningitidis* Z2491.
A;Reference number: A81775; MUID:20222556; PMID:10761919
A;Accession: H81806
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-268 <PAR>
A;Cross-references: UNIPROT:Q9JTB4; GB:AL162757; GB:AL157959; NID:g7380371; PIDN:CAB85038
A;Experimental source: serogroup A, strain Z2491
C;Genetics:
A;Gene: hisJ; NMA1811
C;Superfamily: lysine-arginine-ornithine-binding protein

Query Match 81.8%; Score 18; DB 2; Length 268;
Best Local Similarity 57.1%; Pred. No. 5.5e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
| |||
Db 41 VASNAEF 47

RESULT 55

B81063
amino acid ABC transporter, periplasmic amino acid-binding protein NMB1612 [imported] - *N*
C;Species: *Neisseria meningitidis*
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C;Accession: B81063
R;Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.
Hikey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;
ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Maignani, V.; Pizza, M.
Science 287, 1809-1815, 2000
A;Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ver
A;Title: Complete genome sequence of *Neisseria meningitidis* serogroup B strain MC58.
A;Reference number: A81000; MUID:20175755; PMID:10710307
A;Accession: B81063
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-268 <TET>
A;Cross-references: UNIPROT:Q9JYF0; GB:AE002511; GB:AE002098; NID:g7226857; PIDN:AAF41964
A;Experimental source: serogroup B, strain MC58
C;Genetics:
C;Superfamily: lysine-arginine-ornithine-binding protein

Query Match 81.8%; Score 18; DB 2; Length 268;
Best Local Similarity 57.1%; Pred. No. 5.5e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
| |||
Db 41 VASNAEF 47

RESULT 56

A60253
calretinin - human
N;Alternate names: calbindin 29
C;Species: *Homo sapiens* (man)
C;Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 09-Jul-2004
C;Accession: S14109; A60253
R;Parmentier, M.; Lefort, A.
Eur. J. Biochem. 196, 79-85, 1991
A;Title: Structure of the human brain calcium-binding protein calretinin and its expres
A;Reference number: S14109; MUID:91160569; PMID:2001709

A;Accession: S14109
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-271 <PAR>
A;Cross-references: UNIPROT:P22676; GB:X56667; NID:g29635; PIDN:CAA39991.1; PID:g29636
R;Parmentier, M.
Adv. Exp. Med. Biol. 255, 233-240, 1989
A;Title: The human calbindins: cDNA and gene cloning.
A;Reference number: A60253; MUID:90144185; PMID:2618861
A;Accession: A60253
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-271 <PA2>
C;Genetics:
A;Gene: GDB:CALB2; CAL2
A;Cross-references: GDB:125374; OMIM:114051
A;Map position: 16q22.2-16q22.2
C;Superfamily: calretinin; calmodulin repeat homology
C;Keywords: brain; calcium binding; duplication; EF hand
F;16-48/Domain: calmodulin repeat homology <EF1>
F;63-95/Domain: calmodulin repeat homology <EF2>
F;107-139/Domain: calmodulin repeat homology <EF3>
F;151-183/Domain: calmodulin repeat homology <EF4>
F;195-227/Domain: calmodulin repeat homology <EF5>
F;29,31,33,35,40/Binding site: calcium (Asp, Asp, Asn, Tyr, Glu) #status predicted
F;120,122,124,126,131/Binding site: calcium (Asp, Asp, Ser, Tyr, Glu) #status predicted
F;164,166,168,170,175/Binding site: calcium (Asp, Asn, Asp, Lys, Glu) #status predicted
F;208,210,212,214,219/Binding site: calcium (Asp, Asp, Ser, Tyr, Glu) #status predicted

Query Match 81.8%; Score 18; DB 1; Length 271;
Best Local Similarity 57.1%; Pred. No. 5.6e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VXXAAEF 7
| |||
Db 106 VGSSAEF 112

RESULT 57
S25006
Calretinin - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 09-Jul-2004
C;Accession: S25006; B38396
R;Strauss, K.I.; Jacobowitz, D.M.
submitted to the EMBL Data Library, June 1992
A;Description: Nucleotide sequence of rat calretinin cDNA: Evolutionary conservation of
A;Reference number: S25006
A;Accession: S25006
A;Molecule type: mRNA
A;Residues: 1-271 <STR>
A;Cross-references: UNIPROT:P47728; EMBL:X66974; NID:g55852; PIDN:CAA47385.1; PID:g55853
R;Gabrielides, C.; McCormack, A.L.; Hunt, D.F.; Christakos, S.
Biochemistry 30, 656-662, 1991
A;Title: Brain calbindin-D-28k and an M-r 29 000 calcium binding protein in cerebellum a
A;Reference number: A38396; MUID:91105154; PMID:1988053
A;Accession: B38396
A;Status: preliminary
A;Molecule type: protein
A;Residues: 'E', '3', 'X', '5-6', 'X', '8-10', 'X', '12', 'X', '14-15', 'X', '17-21', 'X', '23', 'X', '42-47', 'X', '49-50',
'8', 'XX', '161', '164', 'XD', '167-170', 'X', '172', 'X', '174-178', 'XX', '181-186', 'XX', '189', 'XX', 'XSVAYK', 'W',
C;Superfamily: calretinin; calmodulin repeat homology
C;Keywords: brain; calcium binding; duplication; EF hand
F;16-48/Domain: calmodulin repeat homology <EF1>
F;63-95/Domain: calmodulin repeat homology <EF2>
F;107-139/Domain: calmodulin repeat homology <EF3>
F;151-183/Domain: calmodulin repeat homology <EF4>
F;195-227/Domain: calmodulin repeat homology <EF5>

Query Match 81.8%; Score 18; DB 1; Length 271;
Best Local Similarity 57.1%; Pred. No. 5.6e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VXXAAEF 7
| |||
Db 106 VGSSAEF 112

RESULT 58
D83914
Chloramphenicol resistance protein BH2116 [imported] - Bacillus halodurans (strain C-125)
C;Species: Bacillus halodurans
C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
A;Accession: D83914
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hirano,
Nucleic Acids Res. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A;Reference number: A83650; MUID:20512582; PMID:11058132
A;Accession: D83914
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-286 <STO>
A;Cross-references: UNIPROT:Q9KB19; GB:AP001514; GB:BA000004; NID:g10174613; PIDN:BA0050;
A;Experimental source: strain C-125
C;Genetics:
A;Gene: BH2116

Query Match 81.8%; Score 18; DB 2; Length 286;
Best Local Similarity 57.1%; Pred. No. 5.9e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VXXAAEF 7
| |||
Db 19 VTGTAEF 25

RESULT 59
H83449
Probable transcription regulator PA1570 [imported] - Pseudomonas aeruginosa (strain PAO1)
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C;Accession: H83449
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Br,
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic patho
A;Reference number: A82950; MUID:20437337; PMID:10984043
A;Accession: H83449
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-293 <STO>
A;Cross-references: UNIPROT:Q9I3E7; GB:AE004585; GB:AE004091; NID:g9947525; PIDN:AAG0495;
A;Experimental source: strain PAO1
C;Genetics:
A;Gene: PA1570
C;Superfamily: probable transcription regulator lsrY

Query Match 81.8%; Score 18; DB 2; Length 293;
Best Local Similarity 57.1%; Pred. No. 6e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VXXAAEF 7
| |||
Db 238 VWSAEF 244

RESULT 60
F84633
Probable carbonyl reductase [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C;Accession: F84633
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.

Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: F84633
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-296 <STO>
A;Cross-references: UNIPROT:Q9ZUHS; GB:AE002093; NID:g4115379; PIDN:AAD03380.1; GSPDB:GN
C;Genetics:
A;Gene: At2g24190
A;Map position: 2

Query Match 81.8%; Score 18; DB 2; Length 296;
Best Local Similarity 57.1%; Pred. No. 6.1e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
|
|
|
Db 74 VTSIAEF 80

RESULT 61
H69312
conserved hypothetical protein AF0504 - *Archaeoglobus fulgidus*
C;Species: *Archaeoglobus fulgidus*
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C;Accession: H69312
R;Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson
; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.
Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.P.; McDonald, L.
Nature 390, 364-370, 1997
A;Authors: Uterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.
Smith, H.O.; Woese, C.R.; Venter, J.C.
A;Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archae
A;Reference number: A69250; MUID:98049343; PMID:9389475
A;Accession: H69312
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-301 <KLE>
A;Cross-references: UNIPROT:O29746; GB:AE001069; GB:AE000782; NID:g2689392; PIDN:AAB9073
C;Superfamily: Methanococcus jannaschii conserved hypothetical protein MJ1629

Query Match 81.8%; Score 18; DB 1; Length 301;
Best Local Similarity 57.1%; Pred. No. 6.2e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
|
|
|
Db 274 VESAAEF 280

RESULT 62
T50027
annexin-like protein - *Arabidopsis thaliana*
N;Alternate names: protein T31P16.220
C;Species: *Arabidopsis thaliana* (mouse-ear cress)
C;Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 09-Jul-2004
C;Accession: T50027
R;Bevan, M.; Zimmermann, W.; Gruenisen, A.; Wamburt, R.; Kalicki, J.; Wohlmann, P.; Sm
submitted to the Protein Sequence Database, May 2000
A;Reference number: 225027
A;Accession: T50027
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-316 <BEV>
A;Cross-references: UNIPROT:Q9LX07; EMBL:AL356332; GSPDB:GN00063; ATSP:T31P16.220
A;Experimental source: cultivar Columbia; BAC clone T31P16
C;Genetics:
A;Gene: ATSP:T31P16.220
A;Map position: 5
A;Introns: 26/1; 74/3; 147/3
C;Superfamily: annexin I; annexin repeat homology

Query Match 81.8%; Score 18; DB 2; Length 316;
Best Local Similarity 57.1%; Pred. No. 6.5e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
|
|
|
Db 268 VTTRAEF 274

RESULT 63
AB4304
hypothetical protein Vng1497c [imported] - *Halobacterium* sp. NRC-1
C;Species: *Halobacterium* sp. NRC-1
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C;Accession: AB4304
R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S
; Leithausen, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablor
Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li
A;Title: Genome sequence of *Halobacterium* species NRC-1.
A;Reference number: A84160; MUID:20504483; PMID:11016950
A;Accession: AB4304
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-317 <STO>
A;Cross-references: UNIPROT:Q9HPS0; GB:AE004437; NID:g10580993; PIDN:AAG19797.1; GSPDB:GN
C;Genetics:
A;Gene: VNG1497C

Query Match 81.8%; Score 18; DB 2; Length 317;
Best Local Similarity 57.1%; Pred. No. 6.5e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
|
|
|
Db 173 VTHAAEF 179

RESULT 64
B72656
hypothetical protein APE0677 - *Aeropyrum pernix* (strain K1)
C;Species: *Aeropyrum pernix*
C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
C;Accession: B72656
R;Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takaha
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Ki
DNA Res. 6, 83-101, 1999
A;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, *Aeropyr*
A;Reference number: A72450; MUID:99310339; PMID:10382966
A;Accession: B72656
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-320 <KAW>
A;Cross-references: UNIPROT:Q9YE96; DDBJ:AP000060; NID:g5104188; PIDN:BAA79650.1; PID:dl
C;Experimental source: strain K1
C;Genetics:
A;Gene: APE0677

Query Match 81.8%; Score 18; DB 2; Length 320;
Best Local Similarity 57.1%; Pred. No. 6.6e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
|
|
|
Db 95 VEAFAEF 101

RESULT 65
C70653
probable prephenate dehydratase - *Mycobacterium tuberculosis* (strain H37RV)
C;Species: *Mycobacterium tuberculosis*
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 17-Feb-2003

A;Experimental source: strain A3(2)
C;Genetics:
A;Gene: SCODB:SC5H1.06c
C;Superfamily: fission yeast pyridoxine 4-dehydrogenase

Query Match 81.8%; Score 18; DB 2; Length 327;
Best Local Similarity 57.1%; Pred. No. 6.9e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
|
DB 248 VEAAAEEF 254

RESULT 68
JCS166
ketol-acid reductoisomerase (EC 1.1.1.86) - Mycobacterium avium
N;Alternate names: acetoxyhydroxy acid isomeroreductase
C;Species: Mycobacterium avium
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C;Accession: JCS166
R;Gusberti, L.; Cantoni, R.; De Rossi, E.; Branzoni, M.; Riccardi, G.
Gene 177, 83-85, 1996
A;Title: Cloning and sequencing of the ilvBNC gene cluster from Mycobacterium avium.
A;Reference number: JCS164; PMID:97080504; PMID:8921849
A;Accession: JCS166
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-333 <GUS>
A;Cross-references: UNIPROT:Q59500; GB:L49392; NID:g1196506; PIDN:AAB38428.1; PID:g1196506
C;Genetics:
A;Gene: ilvC
C;Superfamily: Methanococcus ketol-acid reductoisomerase; ketol-acid reductoisomerase hom
C;Keywords: isoleucine-valine biosynthesis; isomerase; oxidoreductase
F;18-200/Domain: ketol-acid reductoisomerase homology <KAR>

Query Match 81.8%; Score 18; DB 1; Length 333;
Best Local Similarity 57.1%; Pred. No. 6.9e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
|
DB 248 VSDTAEEF 254

RESULT 69
A69059
ketol-acid reductoisomerase - Methanobacterium thermoautotrophicum (strain Delta H)
C;Species: Methanobacterium thermoautotrophicum
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C;Accession: A69059
R;Smith, D.R.; Doucet-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.; F
ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
J. Bacteriol. 179, 7135-7155, 1997
A;Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: functio
A;Reference number: A69000; PMID:98037514; PMID:9371463
A;Accession: A69059
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-333 <MTH>
A;Cross-references: UNIPROT:O27491; GB:AE000905; GB:AE000666; NID:g2622541; PIDN:AAB8591;
A;Experimental source: strain Delta H
C;Genetics:
A;Gene: MTH1442
A;Start codon: TTG
C;Superfamily: Methanococcus ketol-acid reductoisomerase; ketol-acid reductoisomerase hom
F;25-207/Domain: ketol-acid reductoisomerase homology <KAR>

Query Match 81.8%; Score 18; DB 2; Length 333;
Best Local Similarity 57.1%; Pred. No. 6.9e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

A;Residues: 1-338 <KEI>
A;Cross-references: UNIPROT:Q57179; GB:L09232; NID:g551777; PIDN:AAA62431.1; PID:g400336
C;Superfamily: Methanococcus ketol-acid reductoisomerase; ketol-acid reductoisomerase hc
C;Keywords: isoleucine-valine biosynthesis; isomerase; oxidoreductase
F;22-204/Domain: ketol-acid reductoisomerase homology <KAR>

Query Match 81.8%; Score 18; DB 1; Length 338;
Best Local Similarity 57.1%; Pred. No. 7e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXXAEEF 7
| |||
Db 252 VSDTAEEF 258

RESULT 75
C69313
conserved hypothetical protein AF0507 - Archaeoglobus fulgidus
C;Species: Archaeoglobus fulgidus
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C;Accession: C69313
R;Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson
.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.
Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
Nature 390, 364-370, 1997
A;Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Arttach, P.; Kaine, B.P.; Sykes, S.
Smith, H.O.; Woese, C.R.; Ventter, J.C.
A;Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaea
A;Reference number: A69250; MUID:98049343; PMID:9389475
A;Accession: C69313
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-342 <KLE>
A;Cross-references: UNIPROT:O29743; GB:AE001069; GB:AE000782; NID:g2689392; PIDN:AAB9073

Query Match 81.8%; Score 18; DB 2; Length 342;
Best Local Similarity 57.1%; Pred. No. 7e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXXAEEF 7
| |||
Db 153 VATEAEF 159

RESULT 76
T45415
ketol-acid reductoisomerase [imported] - Mycobacterium leprae
C;Species: Mycobacterium leprae
C;Date: 31-Jan-2000 #sequence_revision 31-Jan-2000 #text_change 09-Jun-2000
C;Accession: T45415
R;Parkhill, J.; Barrall, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, September 1997
A;Reference number: Z16918
A;Accession: T45415
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-343 <PAR>
A;Cross-references: EMBL:Z99263; PIDN:CAB16437.1
A;Experimental source: cosmid B637
A;Note: ilvC
C;Genetics:
C;Superfamily: Methanococcus ketol-acid reductoisomerase; ketol-acid reductoisomerase hc
P;28-210/Domain: ketol-acid reductoisomerase homology <KAR>

Query Match 81.8%; Score 18; DB 2; Length 343;
Best Local Similarity 57.1%; Pred. No. 7.1e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXXAEEF 7
| |||
Db 258 VSDTAEEF 264

RESULT 77
B87463
hypothetical protein CCI1726 [imported] - Caulobacter crescentus
C;Species: Caulobacter crescentus
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C;Accession: B87463
R;Nierman, W.C.; Feidiblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A;Title: Complete Genome Sequence of Caulobacter crescentus.
A;Reference number: A87249; MUID:21173698; PMID:11259647
A;Accession: B87463
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-343 <STO>
A;Cross-references: UNIPROT:Q9A7J7; GB:AE005673; NID:gl3423142; PIDN:AAK23702.1; GSPDB:G
C;Genetics:
C;Superfamily: pyruvate dehydrogenase (lipoamide) alpha chain; thiamin pyrophosphate-bin

Query Match 81.8%; Score 18; DB 2; Length 343;
Best Local Similarity 57.1%; Pred. No. 7.1e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXXAEEF 7
| |||
Db 315 VAEAAEF 321

RESULT 78
C82156
conserved hypothetical protein VCI1791 [imported] - Vibrio cholerae (strain N16961 serog
C;Species: Vibrio cholerae
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C;Accession: C82156
R;Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P.
l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A;Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A;Reference number: A82035; MUID:20406833; PMID:10952301
A;Accession: C82156
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-346 <HEI>
A;Cross-references: UNIPROT:Q9KR55; GB:AE004256; GB:AE003852; NID:g9656310; PIDN:AAF9494
A;Experimental source: serogroup O1; strain N16961; biotype El Tor
C;Genetics:
A;Gene: VCI1791
A;Map position: 1

Query Match 81.8%; Score 18; DB 2; Length 346;
Best Local Similarity 57.1%; Pred. No. 7.1e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXXAEEF 7
| |||
Db 76 VKTAAEF 82

RESULT 79
F65131
hypothetical 39.3 kD protein in cysG-trpS intergenic region - Escherichia coli (strain K
C;Species: Escherichia coli
C;Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 01-Mar-2002
C;Accession: F65131
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A;Title: The complete genome sequence of Escherichia coli K-12.
A;Reference number: A64720; MUID:97426617; PMID:9278503
A;Accession: F65131

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA
A;Residues: 1-347 <BLAT>
A;Cross-references: GB:AE000413; GB:U00096; NID:G2367215; PIDN:AAC76396.1; PID:g1789772;
A;Experimental source: strain K-12, substrain MG1655
C;Genetics:
A;Gene: ynfN

Query Match 81.8%; Score 18; DB 2; Length 347;
Best Local Similarity 42.9%; Pred. No. 7.1e+02;
Matches 3; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
: |||
Db 140 ITSAAEF 146

RESULT 80

F91156

probable transport protein ECe4222 [imported] - Escherichia coli (strain O157:H7, substra

C;Species: Escherichia coli

C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004

C;Accession: F91156

R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.

gaseawata, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.

DNA Res. 8, 11-22, 2001

A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gene

A;Reference number: A99629; MUID:21156231; PMID:11258796

A;Accession: F91156

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-347 <HAY>

A;Cross-references: UNIPROT:Q8X844; GB:BA000007; PIDN:BA837645.1; PID:g13363696; GSPDB:G

A;Experimental source: strain O157:H7, substrain RMD 050952

C;Genetics:

A;Gene: ECe4222

Query Match 81.8%; Score 18; DB 2; Length 347;

Best Local Similarity 42.9%; Pred. No. 7.1e+02;

Matches 3; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
: |||
Db 140 ITSAAEF 146

RESULT 81

C86002

probable transport protein ynfN [imported] - Escherichia coli (strain O157:H7, substrain

C;Species: Escherichia coli

C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004

C;Accession: C86002

R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew

iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,

Nature 409, 529-533, 2001

A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.

A;Reference number: A85480; MUID:21074935; PMID:11206551

A;Accession: C86002

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-347 <STO>

A;Cross-references: UNIPROT:Q8X844; GB:AE005174; NID:g12517999; PIDN:AAG58479.1; GSPDB:G

A;Experimental source: strain O157:H7, substrain EDL933

C;Genetics:

A;Gene: ynfN

Query Match 81.8%; Score 18; DB 2; Length 347;
Best Local Similarity 42.9%; Pred. No. 7.1e+02;
Matches 3; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
: |||
Db 140 ITSAAEF 146

RESULT 82

AH1427

phosphoserine aminotransferase homolog serC [imported] - Listeria monocytogenes (strain I

C;Species: Listeria monocytogenes

C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004

C;Accession: AH1427

R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker,

; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Feihi, H.,

D.; Jones, L.M.; Karst, U.

Science 294, 849-852, 2001

A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Mat

ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,

A;Title: Comparative Genomics of Listeria species.

A;Reference number: AB1077; MUID:21537279; PMID:11679669

A;Accession: AH1427

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-363 <GLA>

A;Cross-references: UNIPROT:Q8Y310; GB:NC_003210; PIDN:CAD01038.1; PID:g16412325; GSPDB:G

A;Experimental source: strain EGD-e

C;Genetics:

A;Gene: serC

C;Superfamily: phosphoserine aminotransferase

Query Match 81.8%; Score 18; DB 2; Length 363;
Best Local Similarity 57.1%; Pred. No. 7.5e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
: |||
Db 298 VTNSAEF 304

RESULT 83

AF1801

phosphoserine aminotransferase homolog serC [imported] - Listeria innocua (strain Clip1

C;Species: Listeria innocua

C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004

C;Accession: AF1801

R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker,

; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Feihi, H.,

D.; Jones, L.M.; Karst, U.

Science 294, 849-852, 2001

A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Mat

ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,

A;Title: Comparative Genomics of Listeria species.

A;Reference number: AB1077; MUID:21537279; PMID:11679669

A;Accession: AF1801

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-363 <GLA>

A;Cross-references: UNIPROT:Q926T3; GB:AL592022; PIDN:CAC98182.1; PID:g16415498; GSPDB:G

A;Experimental source: strain Clip11262

C;Genetics:

A;Gene: serC

C;Superfamily: phosphoserine aminotransferase

Query Match 81.8%; Score 18; DB 2; Length 363;
Best Local Similarity 57.1%; Pred. No. 7.5e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
: |||
Db 298 VTNSAEF 304

RESULT 84

H83369

hypothetical protein PA2209 [imported] - Pseudomonas aeruginosa (strain PA01)

C;Species: Pseudomonas aeruginosa

C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004

A;Gene: CESP:C50B6.10
A;Map position: 5
A;Introns: 66/3; 96/1; 122/2; 167/2; 189/3; 210/3; 313/1

RESULT 89
E95261

C;Keywords: carbon-carbon lyase; carboxy-lyase
F;1-405/Product: pyruvate decarboxylase #status experimental <MAT>
P;271-317/Domain: thiamin pyrophosphate-binding domain homology <TPB>

Query Match 81.8%; Score 18; DB 2; Length 405;
Best Local Similarity 57.1%; Pred. No. 8.5e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXXXAEF 7
| |||
DB 41 VEAAXEF 47

RESULT 94

C38351
phosphoprotein phosphatase (EC 3.1.3.16) 2A-beta 55K regulatory chain - rabbit (fragment)
C;Species: Oryctolagus cuniculus (domestic rabbit)
C;Date: 30-Dec-1991 #sequence_revision 30-Dec-1991 #text_change 23-Jun-1993
C;Accession: C38351
R;Mayer, R.E.; Hendrix, P.; Cron, P.; Matthies, R.; Stone, S.R.; Goris, J.; Merlevede, W.
Biochemistry 30, 3589-3597, 1991
A;Title: Structure of the 55-kDa regulatory subunit of protein phosphatase 2A: evidence
A;Reference number: A38351; MUID:91198016; PMID:1849734
A;Accession: C38351
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-412 <MAT>
A;Cross-references: GB:J05328
C;Keywords: phosphoric monoester hydrolase

Query Match 81.8%; Score 18; DB 2; Length 412;
Best Local Similarity 57.1%; Pred. No. 8.5e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXXXAEF 7
| |||
DB 193 VITAAEF 199

RESULT 95

S57820
pyruvate decarboxylase (EC 4.1.1.1) 1 - common tobacco (fragment)
C;Species: Nicotiana tabacum (common tobacco)
C;Date: 28-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 12-Jul-2004
C;Accession: S57820
R;Bucher, M.; Brander, K.A.; Shicego, S.; Mandel, T.; Kuhlmeier, C.
Plant Mol. Biol. 28, 739-750, 1995
A;Title: Aerobic fermentation in tobacco pollen.
A;Reference number: S57819; MUID:95375236; PMID:7647304
A;Accession: S57820
A;Status: preliminary; nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-417 <BUC>
A;Cross-references: EMBL:X81854
C;Superfamily: thiamin pyrophosphate-binding domain homology
C;Keywords: carbon-carbon lyase; carboxy-lyase

Query Match 81.8%; Score 18; DB 2; Length 417;
Best Local Similarity 57.1%; Pred. No. 8.6e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXXXAEF 7
| |||
DB 146 VEAAXEF 152

RESULT 96

D95972
probable sugar uptake ABC transporter periplasmic solute-binding protein precursor SMB21
C;Species: Sinorhizobium meliloti
C;Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004
C;Accession: D95972
R;Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernan

Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001

A;Title: The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-fixing endo-
A;Reference number: A95842; MUID:21396508; PMID:11481431
A;Accession: D95972
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-420 <KUR>
A;Cross-references: UNIPROT:Q926G0; GB:AL591985; PIDN:CAC49444.1; PID:g15140930; GSPDB:G4
A;Experimental source: strain 1021, megaplasmid pSymB
R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubier,
Pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
hebaullt, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
A;Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A;Reference number: A96039; MUID:21368234; PMID:11474104
A;Contents: annotation
C;Genetics:
A;Gene: SMB21604
A;Genome: plasmid

Query Match 81.8%; Score 18; DB 2; Length 420;
Best Local Similarity 57.1%; Pred. No. 8.7e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXXXAEF 7
| |||
DB 350 VTESAEP 356

RESULT 97

S55631
villon protein kinase 36 - equine herpesvirus 2
C;Species: equine herpesvirus 2
C;Date: 27-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 09-Jul-2004
C;Accession: S55631
R;Telford, E.A.R.; Watson, M.S.; Aird, H.C.; Perry, J.; Davison, A.J.
J. Mol. Biol. 249, 520-528, 1995

A;Title: The DNA sequence of equine herpesvirus 2.

A;Reference number: S55594; MUID:95302501; PMID:7783207

A;Accession: S55631

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-438 <TEL>
A;Cross-references: UNIPROT:Q66640; GB:U20824; NID:G695172; PIDN:AAAC13824.1; PID:G695209
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, February 1995

Query Match 81.8%; Score 18; DB 2; Length 438;
Best Local Similarity 57.1%; Pred. No. 9e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXXXAEF 7
| |||
DB 368 VSSGAEP 374

RESULT 98

S65685
protein phosphatase 2A, chain B (BR beta) - rat
C;Species: Rattus sp. (rat)
C;Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 05-Nov-1999
C;Accession: S65685
R;Akiyama, N.; Shima, H.; Hatano, Y.; Osawa, Y.; Sugimura, T.; Nagao, M.
Eur. J. Biochem. 230, 766-772, 1995
A;Title: cDNA cloning of BR-gamma, a novel brain-specific isoform of the B regulatory sub-
A;Reference number: S65685; MUID:95331316; PMID:7607250
A;Accession: S65685
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-443 <AKI>
A;Cross-references: EMBL:D38260; NID:G1065605; PIDN:BAA07412.1; PID:d1007991; PID:g17773

```
Query Match      81.8%; Score 18; DB 2; Length 443;
Best Local Similarity 57.1%; Pred. No. 9.1e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 VXXAAEF 7
      |  |  |  |
Db      224 VITAAEF 230

RESULT 99
B38351
phosphoprotein phosphatase 2-beta regulatory chain - human
N;Alternate names: phosphoprotein phosphatase 2A-beta 55K regulatory chain B
C;Species: Homo sapiens (man)
C;Date: 30-Dec-1991 #sequence_revision 30-Dec-1991 #text_change 09-Jul-2004
C;Accession: B38351
R;Mayer, R.E.; Hendrix, P.; Cron, P.; Matthies, R.; Stone, S.R.; Goris, J.; Merlevede, W
Biochemistry 30, 3589-3597, 1991
A;Title: Structure of the 55-kDa regulatory subunit of protein phosphatase 2A: evidence
A;Reference number: A38351; MUID:91198016; PMID:1849734
A;Accession: B38351
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-443 <MAY>
A;Cross-references: UNIPROT:Q00005; GB:M64930; GB:J05328; NID:gl90423; PIDN:AAA36493.1;
C;Genetics:
A;Gene: GDB:PPP2R2B
A;Cross-references: GDB:136801

Query Match      81.8%; Score 18; DB 2; Length 443;
Best Local Similarity 57.1%; Pred. No. 9.1e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 VXXAAEF 7
      |  |  |  |
Db      224 VITAAEF 230

RESULT 100
A38351
phosphoprotein phosphatase 2-alpha regulatory chain - human
N;Alternate names: phosphoprotein phosphatase 2A-alpha 55K regulatory chain B
C;Species: Homo sapiens (man)
C;Date: 30-Dec-1991 #sequence_revision 30-Dec-1991 #text_change 09-Jul-2004
C;Accession: A38351
R;Mayer, R.E.; Hendrix, P.; Cron, P.; Matthies, R.; Stone, S.R.; Goris, J.; Merlevede, W
Biochemistry 30, 3589-3597, 1991
A;Title: Structure of the 55-kDa regulatory subunit of protein phosphatase 2A: evidence
A;Reference number: A38351; MUID:91198016; PMID:1849734
A;Accession: A38351
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-447 <MAY>
A;Cross-references: UNIPROT:Q00007; GB:M64929; GB:J05328; NID:gl90421; PIDN:AAA36490.1;
C;Genetics:
A;Gene: GDB:PPP2R2A
A;Cross-references: GDB:136800

Query Match      81.8%; Score 18; DB 2; Length 447;
Best Local Similarity 57.1%; Pred. No. 9.2e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 VXXAAEF 7
      |  |  |  |
Db      228 VITAAEF 234
```

Search completed: June 13, 2005, 14:01:42
Job time : 29 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2005, 13:23:36 ; Search time 113 seconds
(without alignments)
31.722 Million cell updates/sec

Title: 09730329-59ED

Perfect score: 22

Sequence: 1 VXXAEF 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 300 summaries

Database :

UniProt_03:*

1: uniprot_sprot:*

2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	19	86.4	116	2 Q8A172	Q8A172 bacteroides
2	19	86.4	126	2 Q7R579	Q7R579 giardia lam
3	19	86.4	146	2 Q8RYJ8	Q8RYJ8 oryza sativ
4	19	86.4	159	2 Q95SM3	Q95SM3 drosophila
5	19	86.4	173	2 Q9NSN9	Q9NSN9 homo sapien
6	19	86.4	222	2 Q6N1D1	Q6N1D1 corynebacte
7	19	86.4	240	2 Q9KU15	Q9KU15 vibrio chol
8	19	86.4	243	2 Q6ZMP3	Q6ZMP3 homo sapien
9	19	86.4	255	2 Q61P2	Q61P2 oryza sativ
10	19	86.4	264	2 Q8X197	Q8X197 clavispora
11	19	86.4	267	2 Q42694	Q42694 candida tro
12	19	86.4	293	2 Q96L14	Q96L14 homo sapien
13	19	86.4	298	1 YFEU_ECOL57	Q8XB12 escherichia
14	19	86.4	298	1 YFEU_ECOL6	Q8FB0 escherichia
15	19	86.4	298	1 YFEU_ECOL1	P76535 escherichia
16	19	86.4	298	1 YFEU_SHIFL	Q83Q4 shigella fl
17	19	86.4	338	2 Q7W036	Q7W036 bordetella
18	19	86.4	355	2 Q6WUC2	Q6WUC2 papaver som
19	19	86.4	377	2 Q72FT8	Q72FT8 desulfobact
20	19	86.4	379	2 Q65KT3	Q65KT3 bacillus li
21	19	86.4	385	2 Q9ZVT5	Q9ZVT5 arabidopsis
22	19	86.4	435	2 Q942T2	Q942T2 oryza sativ
23	19	86.4	462	2 Q88CB9	Q88CB9 pseudomonas
24	19	86.4	484	2 Q85932	Q85932 sphingomonas
25	19	86.4	516	2 Q7MQB5	Q7MQB5 vibrio vuln
26	19	86.4	516	2 Q8DD76	Q8DD76 vibrio vuln
27	19	86.4	522	2 Q6MNS1	Q6MNS1 bdellovibri
28	19	86.4	575	2 Q803T4	Q803T4 brachydanio
29	19	86.4	599	2 Q96Z10	Q96Z10 sulfolobus
30	19	86.4	712	2 Q6C2E4	Q6C2E4 yarrowia li
31	19	86.4	957	2 Q76EM6	Q76EM6 gluconobact

32	19	86.4	977	2 P91775	P91775 pacifastacu
33	19	86.4	995	2 Q89QJ5	Q89QJ5 bradyrhizob
34	19	86.4	1096	2 Q74L34	Q74L34 lactobacill
35	19	86.4	1191	2 Q7QJ34	Q7QJ34 giardia lam
36	19	86.4	1234	2 Q92K44	Q92K44 rhizobium m
37	19	86.4	1266	2 Q823D7	Q823D7 salmonella
38	19	86.4	1266	2 Q8ZLP4	Q8ZLP4 salmonella
39	19	86.4	1266	2 Q83SV7	Q83SV7 salmonella
40	19	86.4	1460	2 Q7LGA9	Q7LGA9 homo sapien
41	19	86.4	1472	2 Q75058	Q75058 homo sapien
42	19	86.4	1486	2 Q9UQ08	Q9UQ08 homo sapien
43	19	86.4	1584	2 Q9UQ09	Q9UQ09 homo sapien
44	19	86.4	1946	2 Q92FH4	Q92FH4 listeria in
45	19	86.4	2106	2 Q6COY6	Q6COY6 yarrowia li
46	19	86.4	2214	2 Q20219	Q20219 caenorhabdi
47	19	86.4	3935	2 Q6E7J8	Q6E7J8 lyngbya maj
48	18	81.8	39	2 Q8F054	Q8F054 leptospira
49	18	81.8	43	2 Q7X3T9	Q7X3T9 xanthomonas
50	18	81.8	76	1 TX3A_AGEAP	P33034 agelenopsis
51	18	81.8	76	1 TX3B_AGEAP	P81744 agelenopsis
52	18	81.8	84	2 Q94CU0	Q94CU0 oryza sativ
53	18	81.8	88	2 Q85470	Q85470 mycobacteri
54	18	81.8	88	2 Q6YVE3	Q6YVE3 oryza sativ
55	18	81.8	92	1 PR10_CAVPO	P20658 cavia porce
56	18	81.8	96	2 Q89WX7	Q89WX7 bradyrhizob
57	18	81.8	102	1 SPT4_KLULA	P81205 kluyveromyc
58	18	81.8	104	2 Q7KOL8	Q7KOL8 plasmodium
59	18	81.8	104	2 Q9NFK9	Q9NFK9 plasmodium
60	18	81.8	106	2 Q8GGE3	Q8GGE3 streptomyce
61	18	81.8	113	2 Q6SK25	Q6SK25 arthroacte
62	18	81.8	121	1 YIF9_YEAST	P40520 saccharomyc
63	18	81.8	123	2 Q8TJ11	Q8TJ11 methanosarc
64	18	81.8	134	2 Q6MK11	Q6MK11 bdellovibri
65	18	81.8	137	2 Q9NZ48	Q9NZ48 homo sapien
66	18	81.8	141	2 Q6N7L1	Q6N7L1 rhodospseudo
67	18	81.8	146	2 Q41050	Q41050 pisum sativ
68	18	81.8	147	2 Q9HNX0	Q9HNX0 halobacteri
69	18	81.8	148	2 Q9CE23	Q9CE23 lactococcus
70	18	81.8	149	1 RL22_PICTO	RL22_PICTO
71	18	81.8	151	2 Q89H55	Q89H55 bradyrhizob
72	18	81.8	159	1 FABZ_CAUCR	Q9A714 caulobacter
73	18	81.8	171	2 Q82RT3	Q82RT3 streptomyce
74	18	81.8	173	2 Q9DDR2	Q9DDR2 xenopus lae
75	18	81.8	179	2 Q8TQ97	Q8TQ97 methanosarc
76	18	81.8	184	2 Q6AC19	Q6AC19 leifsonia x
77	18	81.8	186	2 Q76YI3	Q76YI3 bacterioph
78	18	81.8	189	2 Q9N4I6	Q9N4I6 caenorhabdi
79	18	81.8	193	2 Q6CDT8	Q6CDT8 yarrowia li
80	18	81.8	193	2 Q8ELV1	Q8ELV1 streptococ
81	18	81.8	194	2 Q63EZ7	Q63EZ7 bacillus ce
82	18	81.8	194	2 Q54133	Q54133 streptomyce
83	18	81.8	194	2 Q82KK3	Q82KK3 streptomyce
84	18	81.8	198	2 Q7NXL1	Q7NXL1 chromobacte
85	18	81.8	199	1 RS6_TOBAC	P29345 nicotiana t
86	18	81.8	201	2 Q9NEA9	Q9NEA9 drosophila
87	18	81.8	202	2 Q48517	Q48517 lactobacill
88	18	81.8	207	2 Q6SKE7	Q6SKE7 arthroacte
89	18	81.8	211	2 Q64XJ4	Q64XJ4 bacteroides
90	18	81.8	211	2 Q6V7X4	Q6V7X4 choristoneu
91	18	81.8	212	2 Q6CE44	Q6CE44 yarrowia li
92	18	81.8	212	2 Q9NEB0	Q9NEB0 drosophila
93	18	81.8	212	2 Q9PPH4	Q9PPH4 campylobact
94	18	81.8	212	2 Q97GV8	Q97GV8 clostridium
95	18	81.8	215	2 Q9NE30	Q9NE30 drosophila
96	18	81.8	216	2 Q6N9P6	Q6N9P6 rhodospseudo
97	18	81.8	217	2 Q7PCM7	Q7PCM7 anopheles g
98	18	81.8	217	2 Q8MTV9	Q8MTV9 porphyromon
99	18	81.8	219	2 Q89ZP3	Q89ZP3 bacteroides
100	18	81.8	219	2 Q8FUG5	Q8FUG5 corynebacte
101	18	81.8	220	2 Q893S1	Q893S1 clostridium
102	18	81.8	223	2 Q9NBS0	Q9NBS0 drosophila
103	18	81.8	228	2 Q7M227	Q7M227 pisum sativ
104	18	81.8	228	2 Q7M228	Q7M228 vicia faba

105	18	81.8	229	2	Q75ZQ0	Q75zq0 haloarcula	178	18	81.8	271	2	Q8B3I2	Q8b3i2 human rotav
106	18	81.8	235	2	Q83S19	Q83s19 shigella fl	179	18	81.8	273	2	Q648P1	Q648p1 uncultured
107	18	81.8	239	2	Q6N896	Q6n896 rhodopseudo	180	18	81.8	273	2	Q68H18	Q68h18 porcine rot
108	18	81.8	239	2	Q6D3R4	Q6d3r4 erwinia car	181	18	81.8	274	2	Q68HJ1	Q68hj1 porcine rot
109	18	81.8	240	1	FRP_VIBHA	Q56691 vibrio harv	182	18	81.8	274	2	Q68HJ2	Q68hj2 porcine rot
110	18	81.8	240	1	NFSA_ECOLI	P17117 escherichia	183	18	81.8	275	2	Q8B3H6	Q8b3h6 human rotav
111	18	81.8	240	1	NFSA_SALTY	Q9z522 salmonella	184	18	81.8	275	2	Q8B3H7	Q8b3h7 human rotav
112	18	81.8	240	2	Q8Z848	Q8z848 salmonella	185	18	81.8	275	2	Q8B3H8	Q8b3h8 human rotav
113	18	81.8	240	2	Q7MFU1	Q7mfu1 vibrio vuln	186	18	81.8	275	2	Q8B3I0	Q8b3i0 human rotav
114	18	81.8	240	2	Q7UD78	Q7ud78 shigella fl	187	18	81.8	275	2	Q8B3I1	Q8b3i1 human rotav
115	18	81.8	240	2	Q8D4B5	Q8d4b5 vibrio para	188	18	81.8	278	2	Q91E91	Q91e91 human rotav
116	18	81.8	240	2	Q8D4B5	Q8d4b5 vibrio vuln	189	18	81.8	279	2	Q9EVM9	Q9evm9 pseudomonas
117	18	81.8	240	2	Q8RFJ1	Q8rfj1 escherichia	190	18	81.8	279	2	Q6A990	Q6a990 propionibac
118	18	81.8	240	2	Q8X6S1	Q8x6s1 escherichia	191	18	81.8	279	2	Q617C6	Q617c6 human rotav
119	18	81.8	242	2	Q8CCS7	Q8ccs7 mus musculu	192	18	81.8	280	2	Q91E89	Q91e89 human rotav
120	18	81.8	243	2	Q6LPH5	Q6lph5 photobacter	193	18	81.8	280	2	Q91E89	Q91e89 human rotav
121	18	81.8	243	2	Q71XQ2	Q71xq2 listeria mo	194	18	81.8	282	2	Q6D4G2	Q6d4g2 erwinia car
122	18	81.8	243	2	Q7MTC8	Q7mtc8 porphyromon	195	18	81.8	282	2	Q91E90	Q91e90 human rotav
123	18	81.8	244	2	Q7MG00	Q7mg00 vibrio vuln	196	18	81.8	286	2	Q6PEG1	Q6peg1 homo sapien
124	18	81.8	244	2	Q8D4G7	Q8d4g7 vibrio vuln	197	18	81.8	286	2	Q9XB19	Q9xb19 bacillus ha
125	18	81.8	246	2	Q7TPQ7	Q7tpq7 mus musculu	198	18	81.8	287	1	Y504_ARCFU	Q29746 archaeoglob
126	18	81.8	247	2	Q8A2E1	Q8a2e1 bacterioides	199	18	81.8	288	2	Q6IN90	Q6in90 homo sapien
127	18	81.8	247	2	Q6C334	Q6c334 rotavirus s	200	18	81.8	288	2	Q8DSP0	Q8dsp0 streptococ
128	18	81.8	247	2	Q8V9B1	Q8v9b1 human rotav	201	18	81.8	289	2	Q8B7V0	Q8b7v0 pseudomonas
129	18	81.8	247	2	Q8V9B2	Q8v9b2 human rotav	202	18	81.8	293	2	Q913E7	Q913e7 pseudomonas
130	18	81.8	247	2	Q8V9B3	Q8v9b3 human rotav	203	18	81.8	295	2	Q88AA9	Q88aa9 pseudomonas
131	18	81.8	247	2	Q8V9B4	Q8v9b4 human rotav	204	18	81.8	296	2	Q9ZUH5	Q9zuh5 arabisidopsis
132	18	81.8	247	2	Q8V9B5	Q8v9b5 human rotav	205	18	81.8	302	2	Q9ZRE5	Q9zre5 rhizobium m
133	18	81.8	247	2	Q98653	Q98653 rotavirus s	206	18	81.8	304	2	Q8G974	Q8g974 pseudomonas
134	18	81.8	247	2	Q76QM3	Q76qm3 rotavirus s	207	18	81.8	304	2	Q913G8	Q913g8 pseudomonas
135	18	81.8	247	2	Q76QM4	Q76qm4 rotavirus s	208	18	81.8	304	2	Q889P6	Q889p6 pseudomonas
136	18	81.8	247	2	Q76QM5	Q76qm5 rotavirus s	209	18	81.8	304	2	Q88J23	Q88j23 pseudomonas
137	18	81.8	247	2	Q86160	Q86160 rotavirus s	210	18	81.8	306	2	Q9P5R1	Q9p5r1 neosporea
138	18	81.8	247	2	Q86161	Q86161 rotavirus s	211	18	81.8	309	2	Q9HG05	Q9hg05 penicillium
139	18	81.8	247	2	Q86162	Q86162 rotavirus s	212	18	81.8	311	2	Q73WV3	Q73wv3 mycobacteri
140	18	81.8	247	2	Q86163	Q86163 rotavirus s	213	18	81.8	312	2	Q8P5U8	Q8p5u8 xanthomonas
141	18	81.8	247	2	Q86164	Q86164 rotavirus s	214	18	81.8	312	2	Q8PH69	Q8ph69 xanthomonas
142	18	81.8	247	2	Q86165	Q86165 rotavirus s	215	18	81.8	316	2	Q9C5V2	Q9c5v2 arabisidopsis
143	18	81.8	247	2	Q86166	Q86166 rotavirus s	216	18	81.8	316	2	Q9LX07	Q9lx07 arabisidopsis
144	18	81.8	247	2	Q86167	Q86167 rotavirus s	217	18	81.8	317	2	Q9HP80	Q9hp80 halobacteri
145	18	81.8	247	2	Q86168	Q86168 rotavirus s	218	18	81.8	319	2	Q8P523	Q8p523 xanthomonas
146	18	81.8	247	2	Q86171	Q86171 rotavirus s	219	18	81.8	319	2	Q7NNM6	Q7nnm6 gloeobacter
147	18	81.8	247	2	Q86174	Q86174 rotavirus s	220	18	81.8	320	2	Q9YE96	Q9ye96 aeropyrum p
148	18	81.8	247	2	Q86509	Q86509 rotavirus s	221	18	81.8	321	2	P96240	P96240 mycobacteri
149	18	81.8	247	2	Q86510	Q86510 rotavirus s	222	18	81.8	321	2	Q7TVJ6	Q7tvj6 mycobacteri
150	18	81.8	247	2	Q86512	Q86512 rotavirus s	223	18	81.8	322	2	Q9CDC4	Q9cdc4 mycobacteri
151	18	81.8	249	1	GRPE_SYNY3	Q59978 syntecocyst	224	18	81.8	323	2	Q87UL4	Q87ul4 pseudomonas
152	18	81.8	249	2	Q75C54	Q75c54 ashbya gos	225	18	81.8	324	2	Q9CNE6	Q9cne6 pasteurella
153	18	81.8	249	2	Q9VB53	Q9vb53 drosophila	226	18	81.8	325	2	Q81EW4	Q81ew4 trypanosoma
154	18	81.8	251	2	Q6MNN6	Q6mnn6 bdellovibri	227	18	81.8	327	1	M1AA_XANAC	Q8p1q5 xanthomonas
155	18	81.8	254	2	Q54472	Q54472 serratia ma	228	18	81.8	327	1	M1AA_XANCP	Q8p9x8 xanthomonas
156	18	81.8	254	2	Q9FBX4	Q9fbx4 streptomyc	229	18	81.8	327	2	Q9X7R4	Q9x7r4 streptomyc
157	18	81.8	258	2	Q8NNS8	Q8nns8 corynebacte	230	18	81.8	331	2	Q8XK38	Q8xk38 clostridium
158	18	81.8	258	2	Q68H76	Q68h76 porcine rot	231	18	81.8	332	2	Q6SHK9	Q6shk9 uncultured
159	18	81.8	259	2	Q87819	Q87819 rhizobium l	232	18	81.8	333	1	ILVC_METTH	Q27491 methanobact
160	18	81.8	259	2	Q68H19	Q68h19 porcine rot	233	18	81.8	333	1	ILVC_MYCAY	Q59500 mycobacteri
161	18	81.8	260	1	GLYA_COR51	P50434 corynebacte	234	18	81.8	333	1	ILVC_MYCBO	P65150 mycobacteri
162	18	81.8	260	1	THI4_ARCFU	Q29556 archaeoglob	235	18	81.8	333	1	ILVC_MYCLE	Q33114 mycobacteri
163	18	81.8	260	2	Q65UJ7	Q65uj7 manheimia	236	18	81.8	333	1	ILVC_MYCPA	Q73vh7 mycobacteri
164	18	81.8	261	1	PYRF_YAMOH	P48844 yamadazyma	237	18	81.8	333	2	ILVC_MYCTU	P65149 mycobacteri
165	18	81.8	265	2	Q6UKI5	Q6uk15 rhynchocelia	238	18	81.8	333	2	Q9Y852	Q9y852 colletotric
166	18	81.8	266	2	Q68HJ0	Q68hj0 porcine rot	239	18	81.8	333	2	Q7WT24	Q7wt24 streptomyc
167	18	81.8	268	1	HISJ_NEIGO	Q06758 neisseria g	240	18	81.8	335	1	K6PF_STRPN	Q97rc6 streptococ
168	18	81.8	268	2	Q6XL53	Q6xl53 comamonas t	241	18	81.8	335	1	K6PF_STRR6	Q8dd85 streptococ
169	18	81.8	268	2	Q6XL54	Q6xl54 comamonas t	242	18	81.8	335	2	Q8PPS7	Q8pps7 xanthomonas
170	18	81.8	268	2	Q9JTE4	Q9jte4 neisseria m	243	18	81.8	336	2	Q69NR4	Q69nr4 oryza sativ
171	18	81.8	268	2	Q9JYF0	Q9jyf0 neisseria m	244	18	81.8	337	1	ILVC_COREF	Q8fx1 corynebacte
172	18	81.8	268	2	Q68HJ9	Q68hj9 porcine rot	245	18	81.8	337	2	Q8DGZ5	Q8dgz5 synchococ
173	18	81.8	269	2	Q6PC56	Q6pc56 brachydanio	246	18	81.8	338	1	ILVC_CORGL	Q57179 corynebacte
174	18	81.8	271	1	CLB2_HUMAN	P22676 homo sapien	247	18	81.8	340	1	FLRB_ECOLI	P45540 escherichia
175	18	81.8	271	1	CLB2_MOUSE	Q08331 mus musculu	248	18	81.8	342	2	O29743	Q29743 archaeoglob
176	18	81.8	271	1	CLB2_RAT	P47728 rattus norv	249	18	81.8	342	2	Q951X3	Q951x3 macaca fasc
177	18	81.8	271	1	UPPP_THETN	Q8rb29 thermoanaer	250	18	81.8	343	2	Q9A7J7	Q9a7j7 caulobacter

251 18 81.8 344 2 Q917S2
 252 18 81.8 344 2 Q9CVK2
 253 18 81.8 345 2 Q98MY9
 254 18 81.8 346 2 Q9KR55
 255 18 81.8 347 1 FRLB_EC057
 256 18 81.8 347 2 Q6NMY3
 257 18 81.8 347 2 Q6N130
 258 18 81.8 349 2 Q9ZPS3
 259 18 81.8 355 2 Q9SOP5
 260 18 81.8 357 2 Q8EVL5
 261 18 81.8 359 2 Q94EM8
 262 18 81.8 361 2 Q6ABX6
 263 18 81.8 363 1 SERC_LISIN
 264 18 81.8 363 1 SERC_LISMO
 265 18 81.8 363 2 Q71VT6
 266 18 81.8 365 2 Q6LLD2
 267 18 81.8 365 2 Q911Q8
 268 18 81.8 367 2 Q9CWU3
 269 18 81.8 368 2 Q9K611
 270 18 81.8 371 2 P91984
 271 18 81.8 371 2 Q8NLP4
 272 18 81.8 371 2 Q92QW1
 273 18 81.8 380 2 Q9BVS0
 274 18 81.8 380 2 Q6D4S8
 275 18 81.8 382 2 Q6MMP4
 276 18 81.8 383 2 Q7UZP4
 277 18 81.8 387 2 Q7Z7H3
 278 18 81.8 387 2 Q83B80
 279 18 81.8 388 2 Q7PJ50
 280 18 81.8 388 2 Q8FLY3
 281 18 81.8 391 2 Q9H6P6
 282 18 81.8 391 2 Q9X0T1
 283 18 81.8 392 2 Q97DG9
 284 18 81.8 393 2 Q97N37
 285 18 81.8 396 1 ASSY_HALN1
 286 18 81.8 396 2 Q6UVS4
 287 18 81.8 397 2 Q06670
 288 18 81.8 397 2 Q8DMW2
 289 18 81.8 399 2 Q9GSG9
 290 18 81.8 402 2 Q8DR06
 291 18 81.8 403 2 Q7XSV3
 292 18 81.8 403 2 Q7W717
 293 18 81.8 403 2 Q7W1S4
 294 18 81.8 405 1 DCP2_PEA
 295 18 81.8 405 2 Q9H577
 296 18 81.8 412 1 DAD3_RHILO
 297 18 81.8 413 1 2AB8_RABIT
 298 18 81.8 416 2 Q67N28
 299 18 81.8 418 1 DCP1_TOBAC
 300 18 81.8 419 2 Q6C256

ALIGNMENTS

RESULT 1
 Q8A172 PRELIMINARY; PRT; 116 AA.
 ID Q8A172
 AC Q8A172;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Hypothetical protein.
 GN OrderedLocustNames=BT3794;
 OS Bacteroides thetaiotaomicron.
 OC Bacteria; Bacteroidales; Bacteroides (class); Bacteroidales;
 OC Bacteroidaceae; Bacteroides.
 OX NCBI_TaxID=818;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=VPI-5482 / ATCC 29148;
 RX MEDLINE=22550858; PubMed=12663928; DOI=10.1126/science.1080029;
 RA Xu J., Bjursell M.K., Himrod J., Deng S., Carmichael L.K.,

RA Chiang H.C., Hooper L.V., Gordon J.I.;
 RT "A genomic view of the human-Bacteroides thetaiotaomicron symbiosis.";
 RL Science 299:2074-2076(2003).
 DR EMBL; AE016942; AAO78899.1; --
 KW Complete proteome; Hypothetical protein.
 SQ SEQUENCE 116 AA; 11869 MW; DFAC729A3CB9FAC CRC64;
 Query Match 86.4%; Score 19; DB 2; Length 116;
 Best Local Similarity 57.1%; Pred. No. 1e+03;
 Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 VXXAAEF 7
 Db 110 VAASAEF 116
 RESULT 2
 Q7R579 PRELIMINARY; PRT; 126 AA.
 ID Q7R579
 AC Q7R579;
 DT 01-MAR-2004 (TrEMBLrel. 26, Created)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE GUP_387_95779_96159.
 OS Giardia lamblia ATCC 50803.
 OC Eukaryota; Diplomonadida; Hexamitidae; Giardiinae; Giardia.
 OX NCBI_TaxID=184922;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=WB C6;
 RA Morrison H.G., McArthur A.G., Adam R.D., Aley S.B., Gillin F.D.,
 RA Olsen G.J., Sogin M.L.;
 RT "Draft sequence of the Giardia lamblia genome.";
 RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
 CC -!- CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 DR EMBL; AACB01000006; EAA42501.1; --
 SQ SEQUENCE 126 AA; 13845 MW; 2BE6EEB9F73D7D7D CRC64;
 Query Match 86.4%; Score 19; DB 2; Length 126;
 Best Local Similarity 57.1%; Pred. No. 1e+03;
 Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 VXXAAEF 7
 Db 90 VTTAAEF 96
 RESULT 3
 Q8RYJ8 PRELIMINARY; PRT; 146 AA.
 ID Q8RYJ8
 AC Q8RYJ8;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE B1139B11.11 protein.
 GN Name=B1139B11.11;
 OS Oryza sativa (japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzaceae; Oryza.
 OX NCBI_TaxID=39947;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX PubMed=12447438; DOI=10.1038/nature01184;
 RA Sasaki T., Matsumoto T., Yamamoto K., Sakata K., Baba T., Katayose Y.,
 RA Wu J., Niimura Y., Cheng Z., Nagamura Y., Antonio B.A., Kanamori H.,
 RA Hosokawa S., Maekawa M., Arikawa K., Chiden Y., Hayashi M.,
 RA Okamoto M., Ando T., Aoki H., Arita K., Hamada M., Harada C.,
 RA Hishita S., Honda M., Ichikawa Y., Ikonuma A., Iijima M., Ikeda M.,
 RA Ikono M., Itoh S., Itoh T., Itoh Y., Iwabuchi A., Kamiya K.,
 RA Karasawa W., Katagiri S., Kikuta A., Kobayashi N., Kono I.,

```

RA Machita K., Maehara T., Mizuno H., Mizubayashi T., Mukai Y.,
RA Nagasaki H., Nakashima M., Nakana Y., Nakamichi Y., Nakamura M.,
RA Namiki N., Negishi M., Ohta I., Ono N., Saji S., Sakai, K., Shibata M.,
RA Shimokawa T., Shomura A., Song J., Takazaki Y., Terasawa K., Tsuji K.,
RA Waki K., Yanagata H., Yamane H., Yoshiki S., Yoshihara R., Yukawa K.,
RA Zhong H., Iwama H., Endo T., Ito H., Hahn J.H., Kim H.I., Eun M.Y.,
RA Yano M., Jiang J., Gojobori T.,
RT "The genome sequence and structure of rice chromosome 1.";
RL Nature 420:312-316(2002).
DR EMBL; AP004368; BAB90783.1; -.
DR HSSP; O82040; 1K9U.
DR Gramene; Q8RYJ8; -.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR InterPro; IPR02048; EF-hand.
DR InterPro; IPR010983; EF_Hand_like.
DR Pfam; PF00036; ehand; 3.
DR ProDom; PD000012; EF-hand; 2.
DR SMART; SM00054; EFh; 4.
DR PROSITE; PS00018; EF_HAND; 3.
KW Calcium; Calcium-binding.
SQ SEQUENCE 146 AA; 16257 MW; C6DDE4C8612A2E7C CRC64;

Query Match 86.4%; Score 19; DB 2; Length 146;
Best Local Similarity 57.1%; Pred. No. 1.3e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
Db 2 VAATAEF 8

RESULT 4
Q95SM3 ID Q95SM3 PRELIMINARY; PRT; 159 AA.
AC Q95SM3;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE GH14459p.
GN ORFNames=CG17816;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephyarodea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkely;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA Nunoo J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,
RA Yu C., Lewis S.E., Rubin G.M., Celniker S.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY060702; AAL28250.1; -.
DR FlyBase; FBgn0037525; CG17816.
SQ SEQUENCE 159 AA; 17663 MW; C33175B9B3436C4D CRC64;

Query Match 86.4%; Score 19; DB 2; Length 159;
Best Local Similarity 57.1%; Pred. No. 1.4e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
Db 13 VASAAEF 19

RESULT 5
Q9NSN9 ID Q9NSN9 PRELIMINARY; PRT; 173 AA.
AC Q9NSN9;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)

```

```

DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical protein DKFp564B0982.
GN Name=DKFp564B0982;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Poustka A., Wellenreuther R., Mewes H.W., Weill B., Wiemann S.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL161964; CAB82309.1; -.
KW Hypothetical protein.
SQ SEQUENCE 173 AA; 19157 MW; E3CF034C1E5D4BB CRC64;

Query Match 86.4%; Score 19; DB 2; Length 173;
Best Local Similarity 57.1%; Pred. No. 1.5e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
Db 138 VSAAEF 144

RESULT 6
Q6NIDI ID Q6NIDI PRELIMINARY; PRT; 222 AA.
AC Q6NIDI;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Putative tetR family regulatory protein.
GN Name=amtR; OrderedLocusNames=DIP0846;
OS Corynebacterium diphtheriae.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=1717;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Biotype gravis / NCTC 13129;
RX MEDLINE=22965443; PubMed=14602910; DOI=10.1093/nar/gkg874;
RA Cerdano-Tarraga A.-M., Efstathiou A., Dover L.G., Holden M.T.G.,
RA Pallen M.J., Bentley S.D., Besra G.S., Churcher C.M., James K.D.,
RA De Zoysa A., Chillingworth T., Cronin A., Dowd L., Feltwell T.,
RA Hamlin N., Holroyd S., Jägers K., Moule S., Quail M.A.,
RA Rabinowitsch E., Rutherford K.M., Thomson N.R., Unwin L.,
RA Whitehead S., Barrett B.G., Parkhill J.;
RT "The complete genome sequence and analysis of Corynebacterium
RT diphtheriae NCTC13129."
RL Nucleic Acids Res. 31:6516-6523(2003).
CC -!- SIMILARITY: Contains 1 HTH tetR-type DNA-binding domain.
DR EMBL; BX248356; CAE49362.1; -.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR009057; Homeodomain_like.
DR InterPro; IPR001647; HTH_Tetr.
DR Pfam; PF00440; Tetr_N; 1.
DR PRINTS; PR00455; HTHTETR.
KW Complete proteome; DNA-binding; Transcription;
KW Transcription regulation.
SQ SEQUENCE 222 AA; 24446 MW; 14DD3E055B8C79F1 CRC64;

Query Match 86.4%; Score 19; DB 2; Length 222;
Best Local Similarity 57.1%; Pred. No. 1.9e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
Db 122 VASAEF 128

RESULT 7

```

```

Q9KU15
ID Q9KU15 PRELIMINARY; PRT; 240 AA.
AC Q9KU15;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE NADPH-flavin oxidoreductase.
GN OrderedLocusNames=VC0715;
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=E1 Tor NI6961 / Serotype O1;
RX MEDLINE=20406833; PubMed=10952301; DOI=10.1038/35020000;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.L.,
RA Ermolaeva M.D., Vamathevan J.J., Baas S., Qin H., Dragoi I.,
RA Sellers P., McDonald L.A., Utterback T.R., Fleischmann R.D.,
RA Nierman W.C., White O., Salzberg S.L., Smith H.O., Colwell R.R.,
RA Mekalanos J.J., Venter J.C., Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
cholerae.";
RL Nature 406:477-483(2000).
DR EMBL; AE004158; AAF93880.1; -.
DR PIR; D82288; D82288.
DR HSPR; Q56691; 1BKJ.
DR TIGR; VC0715; -.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR000415; Nitroreductase.
DR Pfam; PF00881; Nitroreductase; 1.
KW Complete proteome.
SQ SEQUENCE 240 AA; 25500 MW; F956E48C3BD6980E CRC64;

Query Match 86.4%; Score 19; DB 2; Length 240;
Best Local Similarity 57.1%; Pred. No. 2e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
DB 70 VASAAEF 76

RESULT 8
Q6ZMP3 PRELIMINARY; PRT; 243 AA.
AC Q6ZMP3;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein FLJ16784.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A.
RC Tissue=Tongue;
RA Tanigami A., Fujiwara T., Shibahara T., Goto Y., Hirao M., Shimizu F.,
RA Wakebe H., Ono T., Hishigaki H., Watanabe T., Ozaki K., Sugiyama T.,
RA Irie R., Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J.,
RA Isono Y., Kawai-Hio Y., Saito K., Nishikawa T., Kimura K.,
RA Yamashita H., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K.,
RA Wagatsuma M., Murakawa K., Kanehori K., Takahashi-Fujii A., Oshima A.,
RA Sugiyama A., Kawakami B., Suzuki Y., Sugano S., Nagahari K.,
RA Maehuo Y., Nagai K., Isogai T.;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK131548; BAD18682.1; -.
DR HSSP; P09215; 1BDY.
DR GO; GO:0016301; F:kinase activity; IEA.

DR GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR InterPro; IPR008973; C2 CalB.
DR InterPro; IPR002219; DAG_PE-bind.
DR Pfam; PF001130; C1_1; 1.
DR PRINTS; PR00008; DAGPEDOMAIN.
DR SMART; SM00109; C1; 1.
DR PROSITE; PS00081; DAG_PE_BIND_DOM_2; 1.
KW Kinase.
SQ SEQUENCE 243 AA; 27566 MW; 23D41825EB9F782D CRC64;

Query Match 86.4%; Score 19; DB 2; Length 243;
Best Local Similarity 57.1%; Pred. No. 2e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
DB 201 VATAAEF 207

RESULT 9
Q69IP2 PRELIMINARY; PRT; 255 AA.
AC Q69IP2;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Ankyrin repeat protein-like.
GN Name=OSUNBA0039D04.6;
OS Oryza sativa (Japonica cultivar-group).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
SEQUENCE FROM N.A.
RA Sasaki T., Matsumoto T., Katayose Y.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 7, BAC
clone:OSUNBA0039D04.";
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP006753; BAD32119.1; -.
DR InterPro; IPR002110; ANK.
DR Pfam; PF00023; Ank; 4.
DR PRINTS; PR01415; ANKYRIN.
DR SMART; SM00248; ANK; 3.
DR PROSITE; PS50088; ANK_REPEAT; 2.
DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
KW ANK repeat.
SQ SEQUENCE 255 AA; 26042 MW; 5E170E5A63808629 CRC64;

Query Match 86.4%; Score 19; DB 2; Length 255;
Best Local Similarity 57.1%; Pred. No. 2.1e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
DB 5 VAAAEF 11

RESULT 10
Q8X197 PRELIMINARY; PRT; 264 AA.
AC Q8X197;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE -Orotidine-5'-phosphate decarboxylase (EC 4.1.1.23).
GN Name=URA3;
OS Clavispora lusitanae.
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Metschnikowiaceae; Clavispora.
OX NCBI_TaxID=36911;
RN [1]
SEQUENCE FROM N.A.

```

```

RC STRAIN=CBS 6936;
RA Francois F., Chapeland-Leclerc F., Villard J., Noel T.;
RT "Development of an integrative transformation system for the
RT opportunistic pathogenic yeast Candida lusitanae using URA3 as a
RT selection marker.";
RL Yeast 21:95-106(2004).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CBS 6936;
RA Francois-Vadrot F., Chapeland-Leclerc F., Villard J., Noel T.;
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
CC -I- CATALYTIC ACTIVITY: Orotidine 5'-phosphate = UMP + CO(2).
CC -I- PATHWAY: Pyrimidine biosynthesis.
CC -I- SIMILARITY: Belongs to the OMP decarboxylase family.
DR EMBL; AF450297; AAL47842.1; -.
DR HSSP; P03962; IDQM.
DR GO; GO:0016829; F:lyase activity; IEA.
DR GO; GO:0004590; F:orotidine-5'-phosphate decarboxylase activity; IEA.
DR GO; GO:0006207; P:'de novo' pyrimidine base biosynthesis; IEA.
DR GO; GO:0006221; P:pyrimidine nucleotide biosynthesis; IEA.
DR Pfam; PF00215; OMPDecase; 1.
DR TIGRFAMs; TIGR01740; PyrF; 1.
DR PROSITE; PS00156; OMPDecase; 1.
KW Decarboxylase; Lyase; Pyrimidine biosynthesis.
SQ SEQUENCE 264 AA; 28913 MW; 98949F050BBEB79B CRC64;

Query Match      86.4%; Score 19; DB 2; Length 264;
Best Local Similarity 57.1%; Pred. No. 2.2e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
DB 37 VSTTAEF 43

RESULT 11
O42694
ID O42694 PRELIMINARY; PRT; 267 AA.
AC O42694;
DT 01-JUN-1998 (TrEMBLrel. 05, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Orotidine-5'-phosphate decarboxylase (EC 4.1.1.23).
GN Name=URA3;
OS Candida tropicalis (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX NCBI_TaxID=5482;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC20336;
RX MEDLINE=98117054; PubMed=9457876;
RA Kanayama N., Ueda M., Atomi H., Tanaka A.;
RT "Genetic evaluation of physiological functions of thiolase isozymes in
RT the n-alkane-assimilating yeast, Candida tropicalis.";
RL J. Bacteriol. 180:690-698(1998).
CC -I- CATALYTIC ACTIVITY: Orotidine 5'-phosphate = UMP + CO(2).
CC -I- PATHWAY: Pyrimidine biosynthesis.
CC -I- SIMILARITY: Belongs to the OMP decarboxylase family.
DR EMBL; AB006207; BAA24611.1; -.
DR HSSP; P03962; IDQM.
DR GO; GO:0016829; F:lyase activity; IEA.
DR GO; GO:0004590; F:orotidine-5'-phosphate decarboxylase activity; IEA.
DR GO; GO:0006207; P:'de novo' pyrimidine base biosynthesis; IEA.
DR GO; GO:0006221; P:pyrimidine nucleotide biosynthesis; IEA.
DR InterPro; IPR01754; OMPDecase.
DR InterPro; IPR011060; Ribp_bind_barrel.
DR Pfam; PF00215; OMPDecase; 1.
DR TIGRFAMs; TIGR01740; PyrF; 1.
DR PROSITE; PS00156; OMPDecase; 1.
KW Decarboxylase; Lyase; Pyrimidine biosynthesis.
SQ SEQUENCE 267 AA; 29521 MW; 48D2C8CBFF95846B CRC64;

```

```

Query Match      86.4%; Score 19; DB 2; Length 267;
Best Local Similarity 57.1%; Pred. No. 2.2e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
DB 39 VTTTAEF 45

RESULT 12
Q96L14
ID Q96L14 PRELIMINARY; PRT; 293 AA.
AC Q96L14;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner K.H., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.P., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skaleka U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Matra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RA Strausberg R.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
RX EMBL; BC014590; AAI14590.2; -.
KW Hypothetical protein.
SQ SEQUENCE 293 AA; 32648 MW; 74904E44B67399F3 CRC64;

Query Match      86.4%; Score 19; DB 2; Length 293;
Best Local Similarity 57.1%; Pred. No. 2.4e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
DB 258 VSNAAEF 264

RESULT 13
YFEU_ECO57
ID YFEU_ECO57 STANDARD; PRT; 298 AA.
AC Q8XB2;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Protein yfeU.
GN Name=yfeU; OrderedLocusNames=z3693, ECs3299;
OS Escherichia coli O157:H7.

```

```
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=83334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / EDL933 / ATCC 700927 / EHEC;
RX MEDLINE=21156231; PubMed=11206551; DOI=10.1038/35054089;
RA Perna N.T., Plunkett G., III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL Nature 409:529-533 (2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / Sakai / RMD 0509952 / EHEC;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Tada T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22 (2001).
CC -1- SIMILARITY: Belongs to the GCKR family.
CC -1- SIMILARITY: Contains 1 SIS domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE005473; BAB357546.1; -.
DR EMBL; AP002561; BAB36722.1; -.
DR PIR; C91041; C91041.1; -.
DR HAMAP; MF_00068; -.
DR InterPro; IPR005486; GCKR.
DR InterPro; IPR005488; GCKR_like.
DR Pfam; PF01380; SIS; 1.
DR TIGRFAMs; TIGR00274; GCKR_like; 1.
DR PROSITE; PS01272; GCKR; 1.
DR Complete proteome. 214
FT DOMAIN 56
SQ SEQUENCE 298 AA; 31112 MW; 231B33A48F7E1B81 CRC64;

Query Match 86.4%; Score 19; DB 1; Length 298;
Best Local Similarity 57.1%; Pred. No. 2.5e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXXAEF 7
DB 172 VSTAEF 178

RESULT 14
YFEU_EC0L6 STANDARD; PRT; 298 AA.
AC Q8FBE0;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Protein yfeU.
GN NamesyfeU; OrderedLocusNames=c2961;
OS Escherichia coli O6.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=217992;
```

```

[1]
SEQUENCE FROM N.A.
STRAIN=O6:H1 / CFT073 / ATCC 700928 / UPEC;
MEDLINE=22388234; PubMed=12471157; DOI=10.1073/pnas.252529799;
Welch R.A., Burland V., Plunkett G., III, Redford P., Roesch P.,
Rasko D.A., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
RT "Extensive mosaic structure revealed by the complete genome sequence
of uropathogenic Escherichia coli.";
Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024 (2002).
CC -1- SIMILARITY: Belongs to the GCKR family.
CC -1- SIMILARITY: Contains 1 SIS domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AS016764; AN81411.1; -.
DR HAMAP; MF_00068; -.
DR InterPro; IPR005486; GCKR.
DR InterPro; IPR005488; GCKR_like.
DR Pfam; PF01380; SIS; 1.
DR TIGRFAMs; TIGR00274; GCKR_like; 1.
DR PROSITE; PS01272; GCKR; 1.
DR Complete proteome. 214
FT DOMAIN 56
SQ SEQUENCE 298 AA; 31105 MW; 3F6DECAE3F784994 CRC64;

Query Match 86.4%; Score 19; DB 1; Length 298;
Best Local Similarity 57.1%; Pred. No. 2.5e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXXAEF 7
DB 172 VSTAEF 178

RESULT 15
YFEU_EC0L1 STANDARD; PRT; 298 AA.
AC P76535; P76965; P76966; P76967;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 25-JAN-2005 (Rel. 46, Last annotation update)
DE Protein yfeU.
GN NamesyfeU; OrderedLocusNames=b2428;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503; DOI=10.1126/science.277.5331.1453;
RA Blattner F.R., Plunkett G., III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
Science 277:1453-1474 (1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=97349980; PubMed=9205837;
RA Yamamoto Y., Aiba H., Baba T., Hayashi K., Inada T., Isono K.,
RA Itoh T., Kimura S., Kitagawa M., Makino K., Miki T., Mitsuhashi N.,
RA Mizobuchi K., Mori H., Nakade S., Nakamura Y., Nashimoto H.,
```

```
RA Oahima T., Oyama S., Saito N., Sampei G., Satoh Y., Sivasundaram S.,
RA Tagami H., Takahashi H., Takeda J., Takemoto K., Uehara K., Wada C.,
RA Yamagata S., Horiuchi T.;
RT "Construction of a contiguous 874-kb sequence of the Escherichia coli-
RT K12 genome corresponding to 50.0-68.8 min on the linkage map and
RT analysis of its sequence features.";
RL DNA Res. 4:91-113(1997).
RN [3]
RP IDENTIFICATION BY MASS SPECTROMETRY.
RX MEDLINE=99420866; PubMed=10493123;
RX DOI=10.1002/ISCI.1522-2683(19990801)20:11<2181::AID-ELPS2181>3.3.CO;2-H;
RA Fountoulakis M., Takacs M.-P., Berndt P., Langer H., Takacs B.;
RT "Enrichment of low abundance proteins of Escherichia coli by
RT hydroxyapatite chromatography.";
RL Electrophoresis 20:2181-2195(1999).
CC -!- SIMILARITY: Belongs to the GCKR family.
CC -!- SIMILARITY: Contains 1 SIS domain.
CC -!- CAUTION: Ref.2 sequence differs from that shown due to a
CC frameshift in position 87.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U00096; AAC75481.1; -
DR EMBL; D90872; BAA16311.1; ALT_FRAME.
DR EMBL; D90872; BAA16312.1; ALT_FRAME.
DR EMBL; D90873; BAA16315.1; -
DR PIR; C65017; C65017.
DR EcoGene; EB3914; -.
DR EcoGene; EG14162; yfeU.
DR HAMAP; MF 00068; -.
DR InterPro; IPR005486; GCKR.
DR InterPro; IPR005488; GCKR_like.
DR Pfam; PF01380; SIS.
DR TIGRFAMs; TIGR00274; GCKR like; 1.
DR PROSITE; PS01272; GCKR; 1.
KW Complete proteome.
FT DOMAIN 56 214 SIS.
SQ SEQUENCE 298 AA; 31220 MW; 34F8F8F78EBB3077 CRC64;

Query Match 86.4%; Score 19; DB 1; Length 298;
Best Local Similarity 57.1%; Pred. No. 2.5e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
Db 172 VSTTAEF 178

RESULT 16
YFEU_SHIFL STANDARD; PRT; 298 AA.
AC Q83QK4;
DT 29-MAR-2004 (Rel. 43, Created)
DT 29-MAR-2004 (Rel. 43, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Protein yfeU.
GN Names=yfeU; OrderedLocusNames=SF2481, S2629;
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=623;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=301 / Serotype 2a;
RX MEDLINE=22272406; PubMed=12384590; DOI=10.1093/nar/gkf566;
RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
```

```
RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
RA Yu J.;
RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
RT through comparison with genomes of Escherichia coli K12 and O157.";
RL Nucleic Acids Res. 30:4432-4441(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=2457T / ATCC 700930 / Serotype 2a;
RX MEDLINE=22590274; PubMed=12704152;
RX DOI=10.1128/JAI.71.5.2775-2786.2003;
RA Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,
RA Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,
RA Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,
RA Schwartz D.C., Blattner F.R.;
RT "Complete genome sequence and comparative genomics of Shigella
RT flexneri serotype 2a strain 2457T.";
RL Infect. Immun. 71:2775-2786(2003).
CC -!- SIMILARITY: Belongs to the GCKR family.
CC -!- SIMILARITY: Contains 1 SIS domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AS015263; AAN43987.1; -
DR EMBL; AS016986; AAP17802.1; -
DR HAMAP; MF 00068; -.
DR InterPro; IPR005486; GCKR.
DR InterPro; IPR005488; GCKR_like.
DR Pfam; PF01380; SIS; 1.
DR TIGRFAMs; TIGR00274; GCKR like; 1.
DR PROSITE; PS01272; GCKR; 1.
KW Complete proteome.
FT DOMAIN 56 214 SIS.
SQ SEQUENCE 298 AA; 31183 MW; FB63F15725C1091F CRC64;

Query Match 86.4%; Score 19; DB 1; Length 298;
Best Local Similarity 57.1%; Pred. No. 2.5e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
Db 172 VSTTAEF 178

RESULT 17
Q7W036 PRELIMINARY; PRT; 338 AA.
AC Q7W036;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=BP0330;
OS Bordetella pertussis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_TaxID=520;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Tohama I / ATCC BAA-589 / NCTC 13251;
RX MEDLINE=22827954; PubMed=12910271; DOI=10.1038/ngl1227;
RA Parkhill J., Sebahia M., Preston A., Murphy L.D., Thomson N.R.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cerdeno-Tarraga A.-M., Temple L., James K.D., Harris B., Quail M.A.,
RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
```

RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
 RA Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,
 RA Leather S., Moule S., Norberczak H., O'Neill S., Ormond D., Price C.,
 RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
 RA Sharp S., Simmonds M., Skelton J., Squares R., Stevens K.,
 RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
 RT "Comparative analysis of the genome sequences of Bordetella pertussis,
 RT Bordetella parapertussis and Bordetella bronchiseptica.",
 RL Nat. Genet. 35:32-40(2003).
 DR EMBL: BX640411; CAB40707.1; -;
 DR GO: GO:0030288; C:periplasmic space (sensu Gram-negative Bact. . .; IEA.
 DR InterPro: IPR005064; UPP0065.
 DR Pfam: PF03401; Bug; 1.
 DR KW Complete proteome; Hypothetical protein.
 SQ SEQUENCE 338 AA; 34618 MW; 65F3D103502CE118 CRC64;
 Query Match 86.4%; Score 19; DB 2; Length 338;
 Best Local Similarity 57.1%; Pred. No. 2.9e+03;
 Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 VXXXXAEF 7
 DB 287 VAASAEF 233
 RESULT 18
 Q6WUC2 PRELIMINARY; PRT; 355 AA.
 ID Q6WUC2
 AC Q6WUC2
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE (R,S)-reticuline 7-O-methyltransferase.
 OS Papaver somniferum (Opium poppy).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; Ranunculales;
 OC Papaveraceae; Papaver.
 OX NCBI_TaxID=3469;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX PubMed=14673446;
 RA Ounatoun A., Decker G., Schmidt J., Lottspeich F., Kutschen T.M.;
 RT "(R,S)-Reticuline 7-O-methyltransferase and (R,S)-norcoclaurine 6-O-
 RT methyltransferase of Papaver somniferum - cDNA cloning and
 RT characterization of methyl transfer enzymes of alkaloid biosynthesis
 RT in opium poppy.",
 RL Plant J. 36:808-819(2003).
 DR EMBL: AY268893; AAQ01668.1; -;
 DR GO: GO:0008171; F:O-methyltransferase activity; IEA.
 DR GO: GO:0008757; F:S-adenosylmethionine-dependent methyltransferase; IEA.
 DR GO: GO:0016740; F:transferase activity; IEA.
 DR InterPro: IPR001601; Methyltransf.
 DR InterPro: IPR000051; SAM_bind.
 DR InterPro: IPR009058; Wing_hlx_DNA_bind.
 KW Methyltransferase; Transferase.
 SQ SEQUENCE 355 AA; 39842 MW; DD408079B18D8F71 CRC64;
 Query Match 86.4%; Score 19; DB 2; Length 355;
 Best Local Similarity 57.1%; Pred. No. 2.9e+03;
 Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 VXXXXAEF 7
 DB 226 VATAAEF 232
 RESULT 19
 Q72FT8 PRELIMINARY; PRT; 377 AA.
 ID Q72FT8
 AC Q72FT8
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)

DE Hypothetical protein.
 GN OrderedLocusNames=DVU0125;
 OS Desulfovibrio vulgaris (strain Hildenborough / ATCC 29579 / NCIMB
 OS 8303).
 OC Bacteria; Proteobacteria; Deltaproteobacteria; Desulfovibrionales;
 OC Desulfovibrionaceae; Desulfovibrio.
 OX NCBI_TaxID=882;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX PubMed=15077118; DOI=10.1038/nbt959;
 RA Heidelberg J.F., Seshadri R., Haveman S.A., Hemme C.L., Paulsen I.T.,
 RA Kolonay J.F., Eisen J.A., Ward N.L., Methe B.A., Brinkac L.M.,
 RA Daugherty S.C., DeBoy R.T., Dodson R.J., Durkin A.S., Madupu R.,
 RA Nelson W.C., Sullivan S.A., Fouts D.E., Haft D.H., Selengut J.,
 RA Peterson J.D., Daviden T.M., Zafar N., Zhou L., Radune D.,
 RA Dimitrov G., Hance M., Tran K., Khouri H.M., Gill J., Utterback T.R.,
 RA Feldblyum T.V., Wall J.D., Voordouw G., Fraser C.M.;
 RT "The genome sequence of the anaerobic, sulfate-reducing bacterium
 RT Desulfovibrio vulgaris Hildenborough.",
 RL Nat. Biotechnol. 22:554-559(2004).
 DR EMBL: AE017309; AAS94609.1; -;
 DR TIGR: DVU0125; -;
 DR InterPro: IPR011053; Hybrid_motif.
 KW Complete proteome.
 SQ SEQUENCE 377 AA; 39322 MW; 51E032F4B7543036 CRC64;
 Query Match 86.4%; Score 19; DB 2; Length 377;
 Best Local Similarity 57.1%; Pred. No. 3.1e+03;
 Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 VXXXXAEF 7
 DB 278 VAASAEF 284
 RESULT 20
 Q65KT3 PRELIMINARY; PRT; 379 AA.
 ID Q65KT3
 AC Q65KT3
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
 DE Hypothetical protein (lambda integrase-like, N-terminal, DNA breaking-
 DE rejoining enzyme, catalytic core).
 GN ORFNames=BL01373, BL01428;
 OS Bacillus licheniformis DSM 13.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=279010;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=DSM 13;
 RX PubMed=15383718;
 RA Veith B., Herzberg C., Steckel S., Feesche J., Maurer K.H.,
 RA Ehrenreich P., Baeumer S., Henne A., Liesegang H., Merkl R.,
 RA Ehrenreich A., Gottschalk G.;
 RT "The Complete Genome Sequence of Bacillus licheniformis DSM13, an
 RT Organism with Great Industrial Potential",
 RL J. Mol. Microbiol. Biotechnol. 7:204-211(2004).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 14580;
 RA Rev M.W., Ramaiya P., Nelson K.A., Brody-Karpin S.D., Zaretsky E.J.,
 RA Tang M., de Leon A.L., Xiang H., Gusti V., Clausen I.G., Olsen P.B.,
 RA Rasmussen M.D., Andersen J.F., Jorgensen P.L., Larsen T.S.,
 RA Sorokin A., Bolotin A., Lapidus A., Gallon N., Ehrlich S.D.,
 RA Berka R.M.;
 RT "Complete genome sequence of the industrial bacterium Bacillus
 RT licheniformis and comparisons with closely related Bacillus species.",
 RL Genome Biol. 5:R77-R77(2004).
 DR EMBL: AE017333; AAU40331.1; -;
 DR EMBL: CP000002; AAU22978.1; -;
 KW Hypothetical protein.
 SQ SEQUENCE 379 AA; 44618 MW; F3D59CE9A9F47F6E CRC64;

```

Query Match      86.4%; Score 19; DB 2; Length 379;
Best Local Similarity 57.1%; Pred. No. 3.1e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VXXAAEF 7
Db 45 VAAAAEF 51

RESULT 21
Q9ZVT5 PRELIMINARY; PRT; 385 AA.
AC Q9ZVT5;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE F15K9.6.
GN Name=F15K9.6;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Vysotskaia V.S., Schwartz J.R., Toriumi M., Yu G., Li J., Liu S.,
RA Kremenetskaia I., Luros J., Araujo R., Buehler E., Conway A.B.,
RA Dwyer K., Feng J., Kim C., Li Y., Shinn P., Sun H., Davis R.W.,
RA Ecker J.R., Federspiel N.A., Theologis A.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC005278; AAC72110.1; -.
DR PIR; A86165; A86165.
SQ SEQUENCE 385 AA; 44026 MW; 37D15C47013040A0A CRC64;

Query Match      86.4%; Score 19; DB 2; Length 385;
Best Local Similarity 57.1%; Pred. No. 3.1e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VXXAAEF 7
Db 172 VSTAAEF 178

RESULT 22
Q942T2 PRELIMINARY; PRT; 435 AA.
AC Q942T2;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE P0583G08.7 protein.
GN Name=P0583G08.7;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA PubMed=12447438; DOI=10.1038/nature01184;
RA Sasaki T., Matsumoto T., Yamamoto K., Sakata K., Baba T., Katayose Y.,
RA Wu J., Nilmura Y., Cheng Z., Nagamura Y., Antonio B.A., Kanamori H.,
RA Hosokawa S., Masukawa M., Arikawa K., Chiden Y., Hayashi M.,
RA Okamoto M., Ando T., Aoki H., Arita K., Hamada M., Harada C.,
RA Hijiwhita S., Honda M., Ichikawa Y., Igonuma A., Iijima M., Ikeda M.,
RA Ikono M., Itoh S., Itoh T., Itoh Y., Itoh Y., Iwabuchi A., Kamiya K.,
RA Karasawa W., Katagiri S., Kikuta A., Kobayashi N., Kono I.,
RA Machita K., Maehara T., Mizuno H., Mizubayashi T., Mukai Y.,
RA Nagaoka H., Nakashima M., Nakama Y., Nakamichi Y., Nakamura M.,
RA Namiki N., Negishi M., Ohta I., Ono N., Saji S., Sakai, K, Shibata M.,
RA Shimokawa T., Shomura A., Song J., Takazaki Y., Terasawa K., Tsuji K.,
RA Waki K., Yamagata H., Yamane H., Yoshiki S., Yoshihara R., Yukawa K.,

```

```

RA Zhong H., Iwama H., Endo T., Ito H., Hahn J.H., Kim H.I., Sun M.Y.,
RA Yano M., Jiang J., Gojobori T.;
RT "The genome sequence and structure of rice chromosome 1.";
RL Nature 420:312-316(2002).
DR EMBL; AP003282; BAB64772.1; -.
DR Gramene; Q942T2; -.
DR InterPro; IPR004314; DUF239.
DR Pfam; PF03080; DUF239; 1.
SQ SEQUENCE 435 AA; 47559 MW; 49CB0F59D6B46DF5 CRC64;

Query Match      86.4%; Score 19; DB 2; Length 435;
Best Local Similarity 57.1%; Pred. No. 3.5e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VXXAAEF 7
Db 213 VATAAEF 219

RESULT 23
Q88CB9 PRELIMINARY; PRT; 462 AA.
AC Q88CB9;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE MATE efflux family protein.
GN OrderedLocustNames=PP5262;
OS Pseudomonas putida (strain KT2440).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=160488;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22423060; PubMed=12534463;
RA Nelson K.E., Weinel C., Paulsen I.T., Dodson R.J., Hilbert H.,
RA Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M.,
RA Brinkac L.M., Beanan M.J., DeBoy R.T., Daugherty S.C., Kolonay J.F.,
RA Madupu R., Nelson W.C., White O., Peterson J.D., Khouri H.M.,
RA Hance I., Chris Lee P., Holtzapple E.K., Scanlan D., Tran K.,
RA Moazzes A., Utterback T.R., Rizzo M., Lee K., Kosack D., Moescl D.,
RA Wedler H., Lauber J., Stjepandic D., Hoheisel J., Straetz M., Heim S.,
RA Kiewitz C., Eisen J.A., Timmis K.N., Duesterhoeft A., Tummeler B.,
RA Fraser C.M.;
RT "Complete genome sequence and comparative analysis of the
RT metabolically versatile Pseudomonas putida KT2440.";
RL Environ. Microbiol. 4:799-808(2002).
DR EMBL; AB016793; AA070827.1; -.
DR TIGR; PP5262; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0015297; F:antiporter activity; IEA.
DR GO; GO:0015238; F:drug transporter activity; IEA.
DR GO; GO:0006855; P:multidrug transport; IEA.
DR InterPro; IPR002528; Mate.
DR Pfam; PF01554; Mate; 2.
DR TIGRFAMs; TIGR00797; mate; 1.
KW Complete proteome.
SQ SEQUENCE 462 AA; 48856 MW; CBDDDA75330F1BA8 CRC64;

Query Match      86.4%; Score 19; DB 2; Length 462;
Best Local Similarity 57.1%; Pred. No. 3.7e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VXXAAEF 7
Db 124 VASAAEF 130

RESULT 24
Q85932 PRELIMINARY; PRT; 484 AA.
AC Q85932;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)

```


DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Pilus assembly and synthesis protein precursor.
GN Namestrah;
OS Sphingomonas aromaticivorans.
OG Plasmid pNLI.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Sphingomonadales;
OC Sphingomonadaceae; Novosphingobium.
OX NCBI_TaxID=48935;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=F199;
RX MEDLINE=99175459; PubMed=10049392;
RA Romine M.F., Stillwell L.C., Wong K.-K., Thurston S.J., Sisk E.C.,
RA Sensen C., Gaasterland T., Fredrickson J.K., Saffer J.D.,
RT "Complete sequence of a 184-kilobase catabolic plasmid from
RT Sphingomonas aromaticivorans F199."
RL J. Bacteriol. 181:1585-1602(1999).
DR EMBL; AF079317; AAD03948.1; -.
DR PIR; T31224; T31224.
DR InterPro; IPR010927; TraH.
DR Pfam; PF06122; TraH; 1.
KW Plasmid; Signal.
KW SIGNAL 1 36 Potential.
FT SEQUENCE 484 AA; 52034 MW; 992DC004E93C4C40 CRC64;
SQ SEQUENCE 484 AA; 52034 MW; 992DC004E93C4C40 CRC64;

Query Match 86.4%; Score 19; DB 2; Length 484;
Best Local Similarity 57.1%; Pred. No. 3.9e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
Db 368 VSAAAEF 374

RESULT 25
Q7MQB5 PRELIMINARY; PRT; 516 AA.
ID Q7MQB5
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Predicted membrane protein.
GN OrderedLocusNames=VV0093;
OS Vibrio vulnificus (strain VJ016).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=196600;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=14656965; DOI=10.1101/gr.1295503;
RA Chen C.-Y., Wu K.-M., Chang Y.-C., Chang C.-H., Tsai H.-C.,
RA Liao T.-L., Liu Y.-M., Chen H.-J., Shen A.B.-T., Li J.-C., Su T.-L.,
RA Shao C.-P., Lee C.-T., Hor L.-I., Tsai S.-F.;
RT "Comparative genome analysis of Vibrio vulnificus, a marine
RT pathogen."
RL Genome Res. 13:2577-2587(2003).
DR EMBL; AP005330; BAC92857.1; -.
KW Complete proteome.
SQ SEQUENCE 516 AA; 57208 MW; 5626021DC1B521A3 CRC64;

Query Match 86.4%; Score 19; DB 2; Length 516;
Best Local Similarity 57.1%; Pred. No. 4.1e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
Db 364 VTSSAEF 370

RESULT 26
Q8DD76 PRELIMINARY; PRT; 516 AA.
ID Q8DD76

AC Q8DD76;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Predicted membrane protein.
GN OrderedLocusNames=VV11134;
OS Vibrio vulnificus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=672;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CMCP6;
RA Rhee J.H., Kim S.Y., Chung S.S., Kim J.J., Moon Y.H., Jeong H.,
RA Choy H.E.;
RT "Complete genome sequence of Vibrio vulnificus CMCP6."
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE016800; AA009608.1; -.
KW Complete proteome.
SQ SEQUENCE 516 AA; 57283 MW; 1DC881279BD2C1EF CRC64;

Query Match 86.4%; Score 19; DB 2; Length 516;
Best Local Similarity 57.1%; Pred. No. 4.1e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
Db 364 VASSAEF 370

RESULT 27
Q6MNS1 PRELIMINARY; PRT; 522 AA.
ID Q6MNS1
AC Q6MNS1;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Microtubule binding protein.
GN OrderedLocusNames=Bdl167;
OS Bdellovibrio bacteriovorus.
OC Bacteria; Proteobacteria; Deltaproteobacteria; Bdellovibrionales;
OC Bdellovibrionaceae; Bdellovibrio.
OX NCBI_TaxID=959;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=14752164; DOI=10.1126/science.1093027;
RX Rendulic S., Jagtap P., Rosinus A., Eppinger M., Baar C., Lanz C.,
RA Keller H., Lambert C., Evans K.J., Goemann A., Meyer F.,
RA Sockett R.E., Schuster S.C.;
RT "A predator unmasked: life cycle of Bdellovibrio bacteriovorus from a
RT genomic perspective."
RL Science 303:689-692(2004).
DR EMBL; BX842649; CAE79080.1; -.
DR GO; GO:0015629; C:actin cytoskeleton; IEA.
DR GO; GO:0003779; P:actin binding; IEA.
DR GO; GO:0007010; P:cytoskeleton organization and biogenesis; IEA.
DR InterPro; IPR002097; Profilin.
KW Complete proteome.
SQ SEQUENCE 522 AA; 58828 MW; 7DCC47F37F143813 CRC64;

Query Match 86.4%; Score 19; DB 2; Length 522;
Best Local Similarity 57.1%; Pred. No. 4.1e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
Db 494 VATAAEF 500

RESULT 28
Q803T4 PRELIMINARY; PRT; 575 AA.
ID Q803T4

```

AC Q803T4;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein zgc:55619.
GN ORFNames=zgc:55619;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AB; TISSUE=Whole body;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner K.H., Schmenk C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Faney J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalek U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=AB; TISSUE=Whole body;
RA Strausberg R.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC044201; AAH44201.1; -.
DR ZFIN; ZDB-GENE-030131-9382; zgc:55619.
KW Hypothetical protein.
SQ SEQUENCE 575 AA; 66162 MW; A3EDACD583A73531 CRC64;

Query Match 86.4%; Score 19; DB 2; Length 575;
Best Local Similarity 57.1%; Pred. No. 4.5e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VVXXAEF 7
Db | |||
17 VATTAEF 23

RESULT 29
Q96210 Q96210 PRELIMINARY; PRT; 599 AA.
AC Q96210;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein ST2017.
GN OrderedLocusNames=ST2017;
OS Sulfolobus tokodaii.
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC Sulfolobus.
OX NCBI_TaxID=111955;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JCM 10545 / 7;
RX MEDLINE=21456156; PubMed=11572479;
RA Kawarabayashi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,
RA Sekine M., Baba S.-I., Ankai A., Kosugi H., Hosoyama A., Fukui S.,

```

```

RA Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,
RA Yoshizawa T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguchi A.,
RA Aoki K.-I., Masuda S., Yanagii M., Nishimura M., Yamagishi A.,
RA Oshima T., Kikuchi H.;
RT "Complete genome sequence of an aerobic thermoacidophilic
RT Crenarchaeon, Sulfolobus tokodaii strain7.";
RL DNA Res. 8:123-140(2001).
DR EMBL; AP000988; BAB67116.1; -.
DR InterPro; IPR011013; Gal_mut_like.
DR InterPro; IPR008928; Glyco_trans_6hp.
DR InterPro; IPR010916; TONB_Box_N.
DR PROSITE; PS00430; TONB_DEPENDENT_REC_1; UNKNOWN_1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 599 AA; 69660 MW; F8B94CAE7731D119 CRC64;

Query Match 86.4%; Score 19; DB 2; Length 599;
Best Local Similarity 57.1%; Pred. No. 4.7e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VVXXAEF 7
Db | |||
67 VTASAEF 73

RESULT 30
Q6C2E4 Q6C2E4 PRELIMINARY; PRT; 712 AA.
ID Q6C2E4;
AC Q6C2E4;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Yarrowia lipolytica chromosome F of strain CLIB99 of Yarrowia
DE lipolytica.
GN ORFNames=YALI0F085919;
OS Yarrowia lipolytica CLIB99.
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Dipodascaceae; Yarrowia.
OX NCBI_TaxID=284591;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CLIB99;
RG Genolevures;
RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,
RA Lafontaine I., de Montigny J., Marck C., Neuvéglise C., Talla E.,
RA Goffard N., Frangeul L., Aigle M., Anchoard V., Babour A., Barbe V.,
RA Barnay S., Blanchin S., Beckerich J.M., Beyne E., Bleykasten C.,
RA Boisrame A., Boyer E., Cattolico L., Confanieri F., de Daruvar A.,
RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,
RA Hantraye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,
RA Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,
RA Nicaud J.M., Nikolski M., Ozias S., Ozier-Kalogeropoulos O.,
RA Pellenz S., Potier S., Richard G.F., Straub M.L., Suleau A.,
RA Swennene D., Tekala F., Wesolowski-Louvel M., Westhof E., Wirth B.,
RA Zeniou-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A.,
RA Bouchier C., Caudron B., Scarpelli C., Gaillardin C., Weissenbach J.,
RA Wincker P., Souciet J.L.;
RT "Genome evolution in yeasts.";
RL Nature 430:35-44(2004).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CLIB99;
RA Genoscope;
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CR382132; CAG77975.1; -.
DR InterPro; IPR007244; Mak10_Actrans.
DR Pfam; PF04112; Mak10; 1.
SQ SEQUENCE 712 AA; 79869 MW; 4E5D512544154E11 CRC64;

Query Match 86.4%; Score 19; DB 2; Length 712;
Best Local Similarity 57.1%; Pred. No. 5.5e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VVXXAEF 7

```

```

Db      300 VTSSAEF 306
RESULT 31
Q76EM6
ID      Q76EM6      PRELIMINARY;      PRT;      957 AA.
AC
DT      05-JUL-2004 (TrEMBLrel. 27, Created)
DT      05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT      05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE      Transaldolase and glucose-6-phosphate isomerase bifunctional
DE      protein.
GN      Name=tal-pgi;
OS      Gluconobacter oxydans (Gluconobacter suboxydans).
OC      Bacteria; Proteobacteria; Alphaproteobacteria; Rhodospirillales;
OC      Acetobacteraceae; Gluconobacter.
OX      NCBI_TaxID=442;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      Sugiyama M., Suzuki S., Tonouchi N., Yokozeki K.;
RL      Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
CC      -1- CATALYTIC ACTIVITY: D-glucose 6-phosphate = D-fructose 6-
CC      phosphate.
CC      -1- PATHWAY: Involved in glycolysis and in gluconeogenesis.
CC      -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC      -1- SIMILARITY: Belongs to the GPI family.
DR      EMBL: AB106333; BAD08593.1; -.
DR      GO: GO:0005737; C:cytoplasm; IEA.
DR      GO: GO:0004347; F:glucose-6-phosphate isomerase activity; IEA.
DR      GO: GO:0016853; F:isomerase activity; IEA.
DR      GO: GO:0004801; F:transaldolase activity; IEA.
DR      GO: GO:0005975; P:carbohydrate metabolism; IEA.
DR      GO: GO:0006094; P:gluconeogenesis; IEA.
DR      GO: GO:0006096; P:glycolysis; IEA.
DR      GO: GO:0006098; P:pentose-phosphate shunt; IEA.
DR      InterPro: IPR001672; GSP_Isomerase.
DR      InterPro: IPR004732; Tal_mycobact.
DR      InterPro: IPR001585; Transaldolase.
DR      Pfam: PF00342; PGI; 1.
DR      Pfam: PF00923; Transaldolase; 1.
DR      PRINTS: PR00662; G6PISOMERASE.
DR      TIGRFAMs: TIGR00876; tal_mycobact; 1.
DR      PROSITE: PS01054; TRANSALDOLASE 1; 1.
KW      Gluconeogenesis; Glycolysis; Isomerase.
SQ      SEQUENCE 957 AA; 103122 MW; 15507DF271E11B07 CRC64;

Query Match      86.4%; Score 19; DB 2; Length 957;
Best Local Similarity 57.1%; Pred. No. 7.2e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 VXXAAEF 7
Db      861 VATAAEF 867
RESULT 32
P91775
ID      P91775      PRELIMINARY;      PRT;      977 AA.
AC
DT      01-MAY-1997 (TrEMBLrel. 03, Created)
DT      01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT      01-MAY-2004 (TrEMBLrel. 26, Last annotation update)
DE      Pacifastin heavy chain precursor.
OS      Pacifastacus leniusculus (Signal crayfish).
OC      Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC      Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Astacidea;
OC      Astacoidea; Astacidae; Pacifastacidae.
OX      NCBI_TaxID=6720;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      Liang Z., Sottrup-Jensen L., Aspan A., Hall M., Soderhall K.;
RL      MEDLINE=97338078; PubMed=9192625; DOI=10.1073/pnas.94.13.6682;
RA      Liang Z., Sottrup-Jensen L., Aspan A., Hall M., Soderhall K.;

Query Match      86.4%; Score 19; DB 2; Length 957;
Best Local Similarity 57.1%; Pred. No. 7.2e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 VXXAAEF 7
Db      861 VATAAEF 867
RESULT 33
Q89QJ5
ID      Q89QJ5      PRELIMINARY;      PRT;      995 AA.
AC
DT      01-JUN-2003 (TrEMBLrel. 24, Created)
DT      01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT      01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE      Glutamate-ammonia-ligase adenylyltransferase.
GN      Name=glnB; OrderedLocusNames=blr3133;
OS      Bradyrhizobium japonicum.
OC      Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC      Bradyrhizobiaceae; Bradyrhizobium.
OX      NCBI_TaxID=375;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiyama T.,
RA      Sasamoto S., Watanabe A., Idesawa K., Iriyuchi M., Kawashima K.,
RA      Kohara M., Matsumoto M., Shimpo S., Tsuruoka H., Wada T., Yamada M.,
RA      Tabata S.;
RT      "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
RT      Bradyrhizobium japonicum USDA110.";
RL      DNA Res. 9:189-197(2002).
DR      EMBL: AP005946; BAC48398.1; -.
DR      GO: GO:0016874; F:ligase activity; IEA.
DR      GO: GO:0016779; F:nucleotidyltransferase activity; IEA.
DR      GO: GO:0016740; F:transferase activity; IEA.
DR      InterPro: IPR005190; GlnE.
DR      Pfam: PF03710; GlnE; 2.
KW      Complete proteome; Ligase; Nucleotidyltransferase; Transferase.
SQ      SEQUENCE 995 AA; 110118 MW; DCF10D729B67A533 CRC64;

Query Match      86.4%; Score 19; DB 2; Length 995;
Best Local Similarity 57.1%; Pred. No. 7.5e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 VXXAAEF 7

```

```
Db      816 VSASAEF 822

RESULT 34
Q74L34      PRELIMINARY;      PRT; 1096 AA.
AC
Q74L34
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=LJ0391;
OS Lactobacillus johnsonii;
OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
OC Lactobacillus.
OX NCBI_TaxID=33959;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCC 533;
RX PubMed=14966310;
RA Priddy R.D., Berger B., Desiere F., Vilanova D., Barretto C.,
RA Pittet A.-C., Zwahlen M.-C., Rouvet M., Altermann E., Barrangou R.,
RA Mollet B., Mercenier A., Klammer T., Arigoni F., Schell M.A.;
RT "The genome sequence of the probiotic intestinal bacterium
RT Lactobacillus johnsonii NCC 533."
RL Proc. Natl. Acad. Sci. U.S.A. 101:2512-2517(2004).
CC -1- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
CC an amide bond (by similarity).
DR EMBL; AE017201; AAS08381.1; -.
DR GO; GO:0009986; C:cell surface; IEA.
DR InterPro; IPR001899; gram_pos_anchor.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR TIGRFAMs; TIGR01167; LPXTG anchor; 1.
DR PROSITE; PS0847; GRAM_POS_ANCHORING; 1.
KW Cell wall; Complete proteome; Hypothetical protein;
KW Peptidoglycan-anchor.
SQ SEQUENCE 1096 AA; 114350 MW; 4C8A938F88FAC284 CRC64;

Query Match      86.4%; Score 19; DB 2; Length 1096;
Best Local Similarity 57.1%; Pred. No. 8.2e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 VXXAAEF 7
Db      289 VSTSAEF 295

RESULT 35
Q7QTJ4      PRELIMINARY;      PRT; 1191 AA.
AC
Q7QTJ4
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE GLP_375_36878_33303.
OS Giardia lamblia ATCC 50803.
OC Eukaryota; Diplomonadida; Hexamitidae; Giardiinae; Giardia.
OX NCBI_TaxID=184922;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WB C6;
RA Morrison H.G., McArthur A.G., Adam R.D., Aley S.B., Gillin F.D.,
RA Olsen G.J., Sogin M.L.;
RA "Draft sequence of the Giardia lamblia genome."
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AACB01000104; EAA38372.1; -.
DR InterPro; IPR002110; ANK.
DR Pfam; PF00023; Ank; 9.
DR PRINTS; PR01415; ANKYRIN.
DR PROSITE; PS50297; ANK_REP_REGION; 1.
```

```
KW ANK repeat.
SQ SEQUENCE 1191 AA; 132435 MW; 70A77664DD550A2D CRC64;

Query Match      86.4%; Score 19; DB 2; Length 1191;
Best Local Similarity 57.1%; Pred. No. 8.8e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 VXXAAEF 7
Db      311 VSATAEF 317

RESULT 36
Q92K44      PRELIMINARY;      PRT; 1234 AA.
AC
Q92K44
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical protein SMC01454.
GN ORFNames=SMC01454;
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1021;
RX MEDLINE=21396507; PubMed=11481430; DOI=10.1073/pnas.161294398;
RA Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,
RA Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,
RA Godrie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masuy D.,
RA Pohl T., Portetelle D., Puehler A., Purnelle B., Ramsperger U.,
RA Renard C., Thebault P., Vandenbol M., Weidner S., Galibert F.;
RT "Analysis of the chromosome sequence of the legume symbiont
RT Sinorhizobium meliloti strain 1021."
RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).
DR EMBL; AL591789; CAC46716.1; -.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 1234 AA; 130531 MW; 772A245CD29F561C CRC64;

Query Match      86.4%; Score 19; DB 2; Length 1234;
Best Local Similarity 57.1%; Pred. No. 9.1e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 VXXAAEF 7
Db      765 VTASAEF 771

RESULT 37
Q8Z3D7      PRELIMINARY;      PRT; 1266 AA.
AC
Q8Z3D7
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Possible exported protein.
GN OrderedLocusNames=STV3549;
OS Salmonella typhi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=601;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CT18;
RX MEDLINE=21534947; PubMed=11677608; DOI=10.1038/35101607;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA Churcher C.M., Mungall K.L., Bentley S.D., Holden M.T.G., Sebaihia M.,
RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,
RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
```

RA Quail M.A., Rutherford K.M., Simmonds M., Skelton J., Stevens K.,
 RA Whitehead S., Barrell B.G.;
 RT "Complete genome sequence of a multiple drug resistant Salmonella
 RT enterica serovar Typhi CT18.";
 RL Nature 413:848-852(2001).
 DR EMBL; AL627278; CAD07884.1; -.
 KW Complete proteome.
 SQ SEQUENCE 1266 AA; 138661 MW; ECEA91FE03086C91 CRC64;

Query Match 86.4%; Score 19; DB 2; Length 1266;
 Best Local Similarity 57.1%; Pred. No. 9.4e+03;
 Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
 | |||
 DB 895 VSSAAEF 901

RESULT 38

Q8ZLP4
 ID Q8ZLP4 PRELIMINARY; PRT; 1266 AA.
 AC Q8ZLP4
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Putative protease.
 GN Names=Vhdp; OrderedLocusNames=STM3369;
 OS Salmonella typhimurium.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Salmonella.
 OX NCBI_TaxID=602;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=LT2;
 RX MEDLINE=21534948; PubMed=11677609; DOI=10.1038/35101614;
 RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
 RA Courtney L., Porwollik S., Ali J., Dante M., Du P., Hou S., Layman D.,
 RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
 RA Ryan E., Sun H., Flores L., Miller W., Stoneking T., Nhan M.,
 RA Waterston R., Wilson R.K.;
 RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
 RT Lt2.";
 RL Nature 413:852-856(2001).
 RL EMBL; AE008855; AAL22238.1; -.
 DR GO; GO:0008233; F:peptidase activity; IEA.
 KW Complete proteome; Protease.
 SQ SEQUENCE 1266 AA; 138756 MW; 455D1BD7ELC687B3 CRC64;

Query Match 86.4%; Score 19; DB 2; Length 1266;
 Best Local Similarity 57.1%; Pred. No. 9.4e+03;
 Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
 | |||
 DB 895 VSSAAEF 901

RESULT 39

Q83SV7
 ID Q83SV7 PRELIMINARY; PRT; 1266 AA.
 AC Q83SV7
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Possible exported protein.
 GN OrderedLocusNames=t3284;
 OS Salmonella typhi.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Salmonella.
 OX NCBI_TaxID=601;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Ty2 / ATCC 700931;

RX MEDLINE=22531367; PubMed=12644504;
 RX DOI=10.1128/JB.185.7.2330-2337.2003;
 RA Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,
 RA Burland V., Kodyonanni V., Schwartz D.C., Blattner F.R.;
 RT "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2
 RT and CT18.";
 RL J. Bacteriol. 185:2330-2337(2003).
 DR EMBL; AE016845; AAO70819.1; -.
 SQ SEQUENCE 1266 AA; 138744 MW; 339C4A88D35922B2 CRC64;

Query Match 86.4%; Score 19; DB 2; Length 1266;
 Best Local Similarity 57.1%; Pred. No. 9.4e+03;
 Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
 | |||
 DB 895 VSSAAEF 901

RESULT 40

Q7LGA9
 ID Q7LGA9 PRELIMINARY; PRT; 1460 AA.
 AC Q7LGA9
 DT 03-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE KARP-1-binding protein 3.
 GN Name=kab;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RA Hara Y., Adachi Y.;
 RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB022659; BAA83380.1; -.
 DR InterPro; IPR000253; FHA.
 DR InterPro; IPR008984; SMAD_FHA.
 DR Pfam; PF00498; FHA; 1.
 DR SMART; SM00240; FHA; 1.
 DR PROSITE; PSS0006; FHA_DOMAIN; 1.
 SQ SEQUENCE 1460 AA; 161436 MW; BAD23EBCA19B65F0 CRC64;

Query Match 86.4%; Score 19; DB 2; Length 1460;
 Best Local Similarity 57.1%; Pred. No. 1.1e+04;
 Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
 | |||
 DB 1425 VSAAEF 1431

RESULT 41

O75058
 ID O75058 PRELIMINARY; PRT; 1472 AA.
 AC O75058
 DT 01-NOV-1998 (TrEMBLrel. 08, Created)
 DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE KIAA0470 protein (Fragment).
 GN Name=KIAA0470;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=98116662; PubMed=9455484;
 RA Seki N., Ohira M., Nagase T., Ishikawa K., Miyajima N., Nakajima D.,
 RA Nomura N., Ohara O.;

```
RT "Characterization of cDNA clones in size-fractionated cDNA libraries
RT from human brain.";
RL DNA Res. 4:345-349(1997).
DR EMBL; AB007939; BAA32315.2; -.
DR PIR; T00095; T00095.
DR InterPro; IPR000253; FHA.
DR InterPro; IPR008984; SMAD_FHA.
DR Pfam; PF00498; FHA; 1.
DR SMART; SM00240; FHA; 1.
DR PROSITE; PSS0006; FHA_DOMAIN; 1.
FT NON TER 1
SQ SEQUENCE 1472 AA; 162839 MW; 81C2CDC135B45093 CRC64;

Query Match 86.4%; Score 19; DB 2; Length 1472;
Best Local Similarity 57.1%; Pred. No. 1.1e+04;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
Db 1437 VSAAAEF 1443

RESULT 42
Q9UQ08 PRELIMINARY; PRT; 1486 AA.
AC Q9UQ08
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE KARP-1-binding protein 2 (KAB2).
GN Name=kab;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Hara Y., Adachi Y.;
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB022658; BAA83379.1; -.
DR InterPro; IPR000253; FHA.
DR InterPro; IPR008984; SMAD_FHA.
DR Pfam; PF00498; FHA; 1.
DR SMART; SM00240; FHA; 1.
DR PROSITE; PSS0006; FHA_DOMAIN; 1.
SQ SEQUENCE 1486 AA; 164568 MW; 79312F1C5CB9D04F CRC64;

Query Match 86.4%; Score 19; DB 2; Length 1486;
Best Local Similarity 57.1%; Pred. No. 1.1e+04;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
Db 1451 VSAAAEF 1457

RESULT 43
Q9UQ09 PRELIMINARY; PRT; 1594 AA.
AC Q9UQ09
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE KARP-1-binding protein 1 (KAB1).
GN Name=kab;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Hara Y., Adachi Y.;
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
RT
```

```
DR EMBL; AB022657; BAA83378.1; -.
DR InterPro; IPR000253; FHA.
DR InterPro; IPR008984; SMAD_FHA.
DR Pfam; PF00498; FHA; 1.
DR SMART; SM00240; FHA; 1.
DR PROSITE; PSS0006; FHA_DOMAIN; 1.
SQ SEQUENCE 1584 AA; 175321 MW; A99D76ED374531F0 CRC64;

Query Match 86.4%; Score 19; DB 2; Length 1584;
Best Local Similarity 57.1%; Pred. No. 1.2e+04;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
Db 1549 VSAAAEF 1555

RESULT 44
Q92FH4 PRELIMINARY; PRT; 1946 AA.
AC Q92FH4
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Lin0132 protein.
GN OrderedLocusNames=lin0132;
OS Listeria innocua.
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OX NCBI_TaxID=1642;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CLIP 11362 / Serovar 6a;
RX MEDLINE=2137279; PubMed=11679669; DOI=10.1126/science.1063447;
RA Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
RA Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA Jones L.-M., Kaerst U., Kretz J., Kuhn M., Kunst F., Kurapkat G.,
RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
RA Nordsiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A.,
RA Vazquez-Boland J.-A., Voss H., Wehlund J., Cossart P.;
RT "Comparative genomics of Listeria species.";
RL Science 294:849-852(2001).
DR EMBL; AL596163; CAC95365.1; -.
DR PIR; AE1449; AE1449.
DR ListiList; LIN0132.
DR InterPro; IPR003961; FN_III.
DR SMART; SM00060; FN3; 1.
DR Complete proteome.
SQ SEQUENCE 1946 AA; 215361 MW; DA1B4B630F9B62B3 CRC64;

Query Match 86.4%; Score 19; DB 2; Length 1946;
Best Local Similarity 57.1%; Pred. No. 1.4e+04;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
Db 928 VSTAAEF 934

RESULT 45
Q6C0Y6 PRELIMINARY; PRT; 2106 AA.
AC Q6C0Y6
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Yarrowia lipolytica chromosome F of strain CLIB99 of Yarrowia
DE lipolytica.
GN ORFNames=YALI0F20724g;
```

OS Yarrowia lipolytica CUB99.
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Dipodascaceae; Yarrowia.
 OX NCBI_TaxID=284591;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CLIB99;
 RG Genolevures;
 RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,
 RA Lafontaine I., de Montigny J., Marck C., Neuveglise C., Talla E.,
 RA Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V.,
 RA Barney A., Blanchin S., Beckerich J.M., Beyne E., Bleykasten C.,
 RA Boisrame A., Boyer J., Cattolico L., Confanioli F., de Daruvar A.,
 RA Despons L., Fabre B., Fairhead C., Ferry-Dumazet H., Groppi A.,
 RA Hantraye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,
 RA Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,
 RA Nicaud J.M., Nikolaki M., Oztas S., Ozier-Kalogeropoulos O.,
 RA Pellenz S., Potier S., Richard G.F., Straub M.L., Suleau A.,
 RA Swennene D., Tekala F., Wesolowski-Louvel M., Westhof E., Wirth B.,
 RA Zeniou-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A.,
 RA Bouchier C., Caudron B., Scarpelli C., Gaillardin C., Weissenbach J.,
 RA Wincker P., Souciet J.L.;
 RT "Genome evolution in yeasts";
 RL Nature 430:35-44 (2004).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CLIB99;
 RG Genoscope;
 RA Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
 RL EMBL; CRJ82132; CAG78485.1; -.
 DR GO; GO:0005634; C:nucleus; IEA.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
 DR GO; GO:0003677; F:DNA binding; IEA.
 DR GO; GO:0003899; F:DNA-directed RNA polymerase activity; IEA.
 DR GO; GO:0000166; F:nucleotide binding; IEA.
 DR GO; GO:0015668; F:type III site-specific deoxyribonuclease ac. .; IEA.
 DR GO; GO:0009307; P:DNA restriction; IEA.
 DR GO; GO:0006350; P:transcription; IEA.
 DR InterPro; IPR003593; AAA_ATPase.
 DR InterPro; IPR001410; DEAD.
 DR InterPro; IPR011545; DEAD/DEAH_N.
 DR InterPro; IPR001650; Helicase_C.
 DR InterPro; IPR006935; ResIII.
 DR InterPro; IPR011260; RNAP_alpha_C.
 DR InterPro; IPR004179; Sec63.
 DR Pfam; PF00270; DEAD; 2.
 DR Pfam; PF00271; Helicase_C; 1.
 DR Pfam; PF04851; ResIII; 1.
 DR Pfam; PF02889; Sec63; 2.
 DR SMART; SM00382; AAA; 2.
 DR SMART; SM00487; DEXDC; 2.
 DR SMART; SM00490; HELIC_C; 1.
 DR SMART; SM00611; SEC63; 2.
 DR SMART; SM00611; SEC63; 2.
 KW ATP-binding; Helicase; Hydrolase.
 SQ SEQUENCE 2106 AA; 234728 MW; 772389CAFP107619 CRC64;

Query Match 86.4%; Score 19; DB 2; Length 2106;
 Best Local Similarity 57.1%; Pred. No. 1.5e+04;
 Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
 |||||
 Db 1810 VTAAEF 1816

RESULT 46

Q20219 PRELIMINARY; PRT; 2214 AA.
 AC Q20219
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

DE Hypothetical protein F40F4.6.
 GN Names=F40F4.6; ORFNames=F40F4.6;
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RX MEDLINE=99069613; PubMed=9851916;
 RG WormBase Consortium;
 RT "Genome sequence of the nematode C. elegans: a platform for
 RT investigating biology. The C. elegans Sequencing Consortium.";
 RL Science 282:2012-2018 (1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Wilson R.;
 RT "The sequence of C. elegans cosmid F40F4.";
 RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Waterston R.;
 RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RG WormBase Consortium;
 RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: Contains 6 EGF-like domains.
 DR EMBL; U04020; AAA81430.1; -.
 DR PIR; T16305; T16305.
 DR WormBase; WEGene00018237; F40F4.6.
 DR WormPep; F40F4.6; CE04536.
 DR GO; GO:0005529; F:sugar binding; IEA.
 DR InterPro; IPR000742; EGF_2.
 DR InterPro; IPR006209; EGF_like.
 DR InterPro; IPR006210; IEGF.
 DR InterPro; IPR001304; Lectin_C.
 DR InterPro; IPR006582; MD.
 DR InterPro; IPR002035; VWF_A.
 DR Pfam; PF00008; EGF; 1.
 DR Pfam; PF00059; Lectin_C; 1.
 DR Pfam; PF00092; VWA; 1.
 DR SMART; SM00034; CLECT; 1.
 DR SMART; SM00181; EGF; 3.
 DR SMART; SM00604; MD; 2.
 DR SMART; SM00327; VWA; 1.
 DR PROSITE; PS50041; C TYPE LECTIN_2; 1.
 DR PROSITE; PS00022; EGF_1; 6.
 DR PROSITE; PS01186; EGF_2; 3.
 DR PROSITE; PS50026; EGF_3; 3.
 DR PROSITE; PS50234; VWFA; 1.
 KW EGF-like domain; Hypothetical protein.
 SQ SEQUENCE 2214 AA; 243435 MW; BOB387B67013A401 CRC64;

Query Match 86.4%; Score 19; DB 2; Length 2214;
 Best Local Similarity 57.1%; Pred. No. 1.6e+04;
 Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
 |||||
 Db 1770 VTSAEF 1776

RESULT 47

Q6E7J8 PRELIMINARY; PRT; 3935 AA.
 ID Q6E7J8
 AC Q6E7J8
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)

```

DE OS Lyngbya majuscula.
OC Bacteria; Cyanobacteria; Oscillatoriales; Lyngbya.
OX NCBI_TaxID=158786;

[1]
RN RP SEQUENCE FROM N.A.
RC STRAIN=JHB;
RX PubMed=15217615; DOI=10.1016/j.chembiol.2004.03.030;
RA Edwards D.J., Marquez B.L., Nogle L.M., McPhail K., Goeger D.E.,
RT "Structure and Biosynthesis of the Jamaicamides, New Mixed Polyketide-
RT peptidic Neurotoxins from the Marine Cyanobacterium Lyngbya
RL majuscula.";
RL Chem. Biol. 11:817-833(2004).
CC -1- COPACTOR: Pyridoxal phosphate (By similarity).
CC -1- SIMILARITY: Belongs to the ATP-dependent AMP-binding enzyme
CC family.
CC -1- SIMILARITY: Belongs to the group II decarboxylase family.
DR EMBL: AY522504; AA598783.1; -.
DR GO: GO:0004024; F:alcohol dehydrogenase activity, zinc-dependent; IEA.
DR GO: GO:0016831; F:carboxyl-lyase activity; IEA.
DR GO: GO:0048037; F:cofactor binding; IEA.
DR GO: GO:0016740; F:transferase activity; IEA.
DR GO: GO:0008270; F:zinc ion binding; IEA.
DR GO: GO:0006520; P:amino acid metabolism; IEA.
DR GO: GO:0006633; P:fatty acid biosynthesis; IEA.
DR GO: GO:0008152; P:metabolism; IEA.
DR InterPro: IPR009081; ACP-like.
DR InterPro: IPR001227; ACP transferase.
DR InterPro: IPR002085; Adh zn family.
DR InterPro: IPR000873; AMP-bind.
DR InterPro: IPR001242; Condensatn.
DR InterPro: IPR011032; GroES-like.
DR InterPro: IPR000794; Ketoacyl_synth.
DR InterPro: IPR006163; Phosphateth_bind.
DR InterPro: IPR006162; Ppantne S.
DR InterPro: IPR002129; Pyridoxal dec.
DR Pfam: PF00698; Acyl_transf_1; 1.
DR Pfam: PF00107; ADH_ZincN_1.
DR Pfam: PF00501; AMP-binding; 1.
DR Pfam: PF00668; Condensation; 1.
DR Pfam: PF00109; ketoacyl-synt; 1.
DR Pfam: PF02801; ketoacyl-synt C; 1.
DR Pfam: PF00550; PP-binding; 2.
DR Pfam: PF00282; Pyridoxal dec; 1.
DR PRINTS: PR00154; AMPBINDING.
DR PROSITE: PS00075; ACP_DOMAIN; 2.
DR PROSITE: PS00606; B_KETOACYL_SYNTHASE; 1.
DR PROSITE: PS00012; PHOSPHOPANTETHEINE; UNKNOWN_1.
KW Lyase; Pyridoxal phosphate; Transferase.
SQ SEQUENCE 3935 AA; 440318 MW; DBF0D4AC0681FF24 CRC64;

Query Match 86.4%; Score 19; DB 2; Length 3935;
Best Local Similarity 57.1%; Pred. No. 2.6e+04;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
| |||
DB 1597 VTAAEF 1603

RESULT 48
R8F054 PRELIMINARY; PRT; 39 AA.
ID Q8F054
AC Q8F054;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein.
GN OrderedLocNames=LA3643;
OS Leptospira interrogans.
OC Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.
OX NCBI_TaxID=173;

```

```

RN RP SEQUENCE FROM N.A.
RC STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar lai;
RX MEDLINE=22598143; PubMed=12712204; DOI=10.1038/nature01597;
RA Ren S.-X., Fu G., Jiang X.-G., Zeng R., Miao Y.-G., Xu H.,
RA Zhang Y.-X., Xiong H., Lu G., Lu L.-F., Jiang H.-Q., Jia J., Tu Y.-F.,
RA Jiang J.-X., Gu W.-Y., Zhang Y.-Q., Cai Z., Sheng H.-H., Yin H.-F.,
RA Zhang Y., Zhu G.-F., Wan M., Huang H.-L., Qian Z., Wang S.-Y., Ma W.,
RA Yao Z.-J., Shen Y., Qiang B.-Q., Xia Q.-C., Guo X.-K., Danchin A.,
RA Saint Girons I., Somerville R.L., Wen Y.-M., Shi M.-H., Chen Z.,
RA Xu J.-G., Zhao G.-P.;
RT "Unique physiological and pathogenic features of Leptospira
RT interrogans revealed by whole-genome sequencing.";
RL Nature 422:888-893(2003).
DR EMBL: AB011520; AAN50841.1; -.
KW Complete proteome.
SQ SEQUENCE 39 AA; 4743 MW; AF11ED6D2C89107B CRC64;

Query Match 81.8%; Score 18; DB 2; Length 39;
Best Local Similarity 57.1%; Pred. No. 7.6e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
| |||
DB 9 VGTAAEF 15

RESULT 49
Q7X3T9 PRELIMINARY; PRT; 43 AA.
ID Q7X3T9
AC Q7X3T9;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Pilus biogenesis protein (Fragment).
GN Name=pilB;
OS Xanthomonas axonopodis (pv. citri).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=92829;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=XW47;
RX PubMed=15057449;
RA Yang Y.C., Chou C.P., Kuo T.T., Lin S.H., Yang M.K.;
RT "PilR enhances the sensitivity of Xanthomonas axonopodis pv. citri to
RT the infection of filamentous bacteriophage Cf.";
RL Curr. Microbiol. 48:251-261(2004).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=XW47;
RA Kou T.T., Yang Y.C., Chou C.P., Yang M.K.;
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY295878; AAP55204.1; -.
FT NON_TER 1 1
SQ SEQUENCE 43 AA; 4574 MW; D0AF93CEDF8DA62 CRC64;

Query Match 81.8%; Score 18; DB 2; Length 43;
Best Local Similarity 57.1%; Pred. No. 8.3e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
| |||
DB 15 VTSLAEF 21

RESULT 50
TX3A_AGEAP
ID TX3A_AGEAP STANDARD; PRT; 76 AA.
AC F33034;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)

```


Query Match 81.8%; Score 18; DB 1; Length 76;

```

RX MEDLINE=22592660; PubMed=12705866; DOI=10.1016/S0092-8674(03)00233-2;
RA Pedulla M.L., Ford M.E., Houtz J.M., Karthikeyan T., Wadsworth C.,
RA Lewis J.A., Jacobs-Sera D., Falbo J., Gross J., Pannunzio N.R.,
RA Brucker W., Kumar V., Kandasamy J., Keenan L., Bardarov S.,
RA Kriakov J., Lawrence J.G., Jacobs W.R. Jr., Hendrix R.W.,
RA Hatfull G.F.;
RT "Origins of highly mosaic mycobacteriophage genomes.";
RL Cell 113.171-182(2003).
DR EMBL; AY129338; AANI2718.1; -. 5DL3C1010CA6985E CRC64;
SQ SEQUENCE 88 AA; 10324 MW; 5DL3C1010CA6985E CRC64;

Query Match 81.8%; Score 18; DB 2; Length 88;
Best Local Similarity 57.1%; Pred. No. 1.6e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
DB 50 VSAAEF 56

RESULT 54
Q6YVE3 PRELIMINARY; PRT; 88 AA.
AC Q6YVE3;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein B1120F06.135 (Hypothetical protein
DE B1272H04.1).
GN Name=B1120F06.135; Synonyms=B1272H04.1;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN 1;
RP SEQUENCE FROM N.A.
RA Sasaki T., Matsumoto T., Katayose Y.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 7, BAC
RT clone:B1120F06.";
RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
RN 2;
RP SEQUENCE FROM N.A.
RA Sasaki T., Matsumoto T., Katayose Y.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 7, BAC
RT clone:B1272H04.";
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP005830; BAC84533.1; -.
DR EMBL; AP006479; BAD32057.1; -.
KW Hypothetical protein.
SQ SEQUENCE 88 AA; 10001 MW; 693B4D3224B4D010 CRC64;

Query Match 81.8%; Score 18; DB 2; Length 88;
Best Local Similarity 57.1%; Pred. No. 1.6e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
DB 52 VATAEF 58

RESULT 55
PR10_CAVPO STANDARD; PRT; 92 AA.
AC P20658;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 05-JUN-2004 (Rel. 44, Last annotation update)
DE Protein 10 (29 kDa calcium-binding protein, brain-specific)
DE (Fragments).
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.

```

```

OX NCBI_TaxID=10141;
RN 1;
RP SEQUENCE.
RC TISSUE=Brain;
RX MEDLINE=90099311; PubMed=2602362;
RA Winsky L., Nakata H., Martin B.M., Jacobowitz D.M.;
RT "Isolation, partial amino acid sequence, and immunohistochemical
RT localization of a brain-specific calcium-binding protein.";
RL Proc. Natl. Acad. Sci. U.S.A. 86:10139-10143(1989).
CC -!- TISSUE SPECIFICITY: Brain.
CC -!- MISCELLANEOUS: Binds calcium.
CC -!- SIMILARITY: Belongs to the calbindin family.
CC -!- SIMILARITY: Contains 6 EF-hand calcium-binding domains.
DR PIR; A34520; A34520.
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR010983; EF-hand like.
DR PROSITE; PS00018; EF_HAND; PARTIAL.
KW Calcium-binding; Direct protein sequencing.
FT NON_TER 1
FT NON_CONS 17 18
FT NON_CONS 27 28
FT NON_CONS 41 42
FT NON_CONS 76 77
FT NON_CONS 82 83
FT NON_TER 92 92
SQ SEQUENCE 92 AA; 11067 MW; D1CDD66C148A7F6A CRC64;

Query Match 81.8%; Score 18; DB 1; Length 92;
Best Local Similarity 57.1%; Pred. No. 1.7e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
DB 30 VGSSAEF 36

RESULT 56
Q89WX7 PRELIMINARY; PRT; 96 AA.
AC Q89WX7;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Bal0551 protein.
GN OrderedLocustNames=bsl0551;
OS Bradyrhizobium japonicum.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Bradyrhizobium.
OX NCBI_TaxID=375;
RN 1;
RP SEQUENCE FROM N.A.
RC STRAIN=USDA110;
RX MEDLINE=22484998; PubMed=12597275;
RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiumi T.,
RA Sasanoto S., Watanabe A., Idegawa K., Iriguchi M., Kawashima K.,
RA Kohara M., Matsumoto M., Shimpo S., Tsuruoka H., Wada T., Yamada M.,
RA Tabata S.;
RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
RT Bradyrhizobium japonicum USDA110.";
RL DNA Res. 9:189-197(2002).
DR EMBL; AP005936; BAC45816.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR003425; Unk_YGGT.
DR Pfam; PF02325; YGGT; 1.
KW Complete proteome.
SQ SEQUENCE 96 AA; 10976 MW; A9E11F9AA4BC7734 CRC64;

Query Match 81.8%; Score 18; DB 2; Length 96;
Best Local Similarity 57.1%; Pred. No. 1.7e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
DB 30 VGSSAEF 36

```

```

Db          41 VSAVAEF 47
RESULT 57
SPT4_KLUJA
ID -SPT4_KLUJA STANDARD; PRT; 102 AA.
AC P81205;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Transcription initiation protein SPT4.
GN Name=SPT4; OrderedLocusNames=KLUJA0807997g;
OS Kluveromyces lactis (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Kluveromyces.
OX NCBI_TaxID=28985;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JBD100;
RX MEDLINE=99087905; PubMed=9871119;
RA Hikkel I., Gbelska Y., Subik J.;
RT "Identification and functional analysis of a Kluveromyces lactis
RT homologue of the SPT4 gene of Saccharomyces cerevisiae.";
RL Curr. Genet. 34:375-378 (1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CBS 2359 / IFO 1267 / NRRL Y-1140 / WM37;
RX PubMed=15229592; DOI=10.1038/nature02579;
RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,
RA Lefontaine I., de Montigny J., Marck C., Neugeglise C., Talla E.,
RA Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V.,
RA Barnay S., Blanchin S., Beckerich J.-M., Beyne E., Bleykasten C.,
RA Boistame A., Boyer J., Cattolico L., Confanioleri F., de Daruvar A.,
RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,
RA Hantraye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,
RA Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,
RA Nicaud J.-M., Nikolski M., Ostas S., Ozier-Kalogoropoulos O.,
RA Pellens S., Potier S., Richard G.F., Straub M.-L., Suleau A.,
RA Swennene D., Tekai F., Wesolowski-Louvel M., Westhof E., Wirth B.,
RA Zeniou-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A.,
RA Bouchier C., Caudron B., Scarpelli C., Gaillardin C., Weissenbach J.,
RA Wincker P., Souciet J.-L.;
RT "Genome evolution in yeasts.";
RL Nature 430:35-44 (2004).
CC -I- FUNCTION: Important for transcription initiation. The SPT4/5/6
CC complex is required for transcription initiation. It may normally
CC act to repress transcription at a variety of loci, and also plays
CC a role in chromatin structure or assembly. Metal binding is
CC important for its function (By similarity).
CC -I- SUBUNIT: Complex of SPT4, SPT5 and SPT6 (By similarity).
CC -I- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; CR382122; CAH02279.1; -
DR InterPro; IPR009287; Spt4.
DR Pfam; PF06093; Spt4; 1.
KW Nuclear protein; Transcription; Zinc-finger.
FT ZN FING 27 By similarity.
SQ SEQUENCE 102 AA; 11224 MW; 869070285CF7917C CRC64;

Query Match 81.8%; Score 18; DB 1; Length 102;
Best Local Similarity 57.1%; Pred. No. 1.8e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
| |||

Db          41 VSAVAEF 47
RESULT 58
SPT4_KLUJA
ID -SPT4_KLUJA PRELIMINARY; PRT; 104 AA.
AC P81205;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Thioredoxin.
GN ORFNames=PF14_0545;
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22255705; PubMed=12368864; DOI=10.1038/nature01097;
RA Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,
RA Carlton J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,
RA Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
RA Chan M.S., Nene V., Shallow S.J., Suh B., Peterson J., Angiuoli S.,
RA Portea M., Allen J., Selengut J., Haft D., Mather M.W., Vaidya A.B.,
RA Martin D.M., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
RA McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,
RA Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
RA Fraser C.M., Barrell B.;
RT "Genome sequence of the human malaria parasite Plasmodium
RT falciparum.";
RL Nature 419:498-511 (2002).
DR EMBL; AE014825; AAN37158.1; -.
DR HSSP; P10599; 1AUC.
DR GO; GO:0005489; F:electron transporter activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR006662; Thioredoxin.
DR InterPro; IPR005746; Thioredoxin.
DR Pfam; PF00085; Thioredoxin; 1.
DR PRINTS; PR00421; THIOREDOXIN.
DR TIGRFAme; TIGR01068; thioredoxin; 1.
DR PROSITE; PS00194; THIOREDOXIN; 1.
KW Redox-active center.
SQ SEQUENCE 104 AA; 11716 MW; 0129998AEB88770C CRC64;

Query Match 81.8%; Score 18; DB 2; Length 104;
Best Local Similarity 57.1%; Pred. No. 1.9e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
| |||

Db          41 VTSQAEF 11
RESULT 59
ID Q9NFK9 PRELIMINARY; PRT; 104 AA.
AC Q9NFK9;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Thioredoxin.
GN Name=trx;
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=3D7;
RA Krnjajski Z., Gilberger T.W., Walter R.D., Mueller S.;
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ277839; CAB90828.1; -.
DR HSSP; P80028; 1EP7.
DR GO; GO:0005489; F:electron transporter activity; IEA.

```

```

DR GO: 0006118; P: electron transport; IEA.
DR InterPro: IPR006662; Thioired.
DR InterPro: IPR005746; Thioiredoxin.
DR InterPro: IPR006663; Thioiredox_dom2.
DR Pfam: PF00085; Thioiredoxin; 1.
DR PRINTS: PR00421; THIOREDOXIN.
DR TIGRFAMs: TIGR01068; thioiredoxin; 1.
DR PROSITE: PS00194; THIOREDOXIN; 1.
KW Redox-active center.
SQ SEQUENCE 104 AA; 11716 MW; 0129998AEB88770C CRC64;

Query Match      81.8%; Score 18; DB 2; Length 104;
Best Local Similarity 57.1%; Pred. No. 1.9e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
   | |||
Db 5 VTSQAEF 11

RESULT 60
Q8GGE3 PRELIMINARY; PRT; 106 AA.
AC Q8GGE3;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein.
OS Streptomyces sp. EN27.
CG Plasmid pEN2701.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=211464;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EN27;
RX MEDLINE=22472596; PubMed=12584005; DOI=10.1016/S0147-619X(02)00153-1;
RA Coombs J.T., Franco C.M.M., Loria R.;
RT "Complete sequencing and analysis of pEN2701, a novel 13-kb plasmid
from an endophytic Streptomyces sp.";
RL Plasmid 49:86-92(2003).
DR EMBL: AF533985; AAN76283.1; -.
KW Hypothetical protein; Plasmid.
SQ SEQUENCE 106 AA; 11488 MW; E2B127026C785A91 CRC64;

Query Match      81.8%; Score 18; DB 2; Length 106;
Best Local Similarity 57.1%; Pred. No. 1.9e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
   | |||
Db 94 VAAQAEF 100

RESULT 61
Q6SK25 PRELIMINARY; PRT; 113 AA.
AC Q6SK25;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
OS Arthrobacter aureus.
OG Plasmid pAA1.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Micrococineae; Micrococcaceae; Arthrobacter.
OX NCBI_TaxID=43663;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TC1;
RX PubMed=15240330; DOI=10.1128/AEM.70.7.4402-4407.2004;
RA Sajjaphan K., Shapir K., Wackett L.P., Palmer M., Blackmon B.,
RA Tomkins J., Sadowsky M.J.;

DR GO: 0006118; P: electron transport; IEA.
DR InterPro: IPR006662; Thioired.
DR InterPro: IPR005746; Thioiredoxin.
DR InterPro: IPR006663; Thioiredox_dom2.
DR Pfam: PF00085; Thioiredoxin; 1.
DR PRINTS: PR00421; THIOREDOXIN.
DR TIGRFAMs: TIGR01068; thioiredoxin; 1.
DR PROSITE: PS00194; THIOREDOXIN; 1.
KW Redox-active center.
SQ SEQUENCE 104 AA; 11716 MW; 0129998AEB88770C CRC64;

Query Match      81.8%; Score 18; DB 2; Length 104;
Best Local Similarity 57.1%; Pred. No. 1.9e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
   | |||
Db 5 VTSQAEF 11

RESULT 60
Q8GGE3 PRELIMINARY; PRT; 106 AA.
AC Q8GGE3;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein.
OS Streptomyces sp. EN27.
CG Plasmid pEN2701.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=211464;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EN27;
RX MEDLINE=22472596; PubMed=12584005; DOI=10.1016/S0147-619X(02)00153-1;
RA Coombs J.T., Franco C.M.M., Loria R.;
RT "Complete sequencing and analysis of pEN2701, a novel 13-kb plasmid
from an endophytic Streptomyces sp.";
RL Plasmid 49:86-92(2003).
DR EMBL: AF533985; AAN76283.1; -.
KW Hypothetical protein; Plasmid.
SQ SEQUENCE 106 AA; 11488 MW; E2B127026C785A91 CRC64;

Query Match      81.8%; Score 18; DB 2; Length 106;
Best Local Similarity 57.1%; Pred. No. 1.9e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
   | |||
Db 94 VAAQAEF 100

RESULT 61
Q6SK25 PRELIMINARY; PRT; 113 AA.
AC Q6SK25;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
OS Arthrobacter aureus.
OG Plasmid pAA1.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Micrococineae; Micrococcaceae; Arthrobacter.
OX NCBI_TaxID=43663;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TC1;
RX PubMed=15240330; DOI=10.1128/AEM.70.7.4402-4407.2004;
RA Sajjaphan K., Shapir K., Wackett L.P., Palmer M., Blackmon B.,
RA Tomkins J., Sadowsky M.J.;

"Arthrobacter aureusens TC1 Atrazine Catabolism Genes trzN, atzB, and
atzC Are Linked on a 160-Kilobase Region and Are Functional in
RT Escherichia coli.";
RL Appl. Environ. Microbiol. 70:4402-4407(2004).
DR EMBL: AY456696; AAS20147.1; -.
DR InterPro: IPR008651; HicB.
DR InterPro: IPR010985; Met_repress_like.
DR Pfam: PF05534; HicB; 1.
KW Hypothetical protein; Plasmid.
SQ SEQUENCE 113 AA; 12762 MW; 85AF588567AAS328 CRC64;

Query Match      81.8%; Score 18; DB 2; Length 113;
Best Local Similarity 57.1%; Pred. No. 2e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
   | |||
Db 26 VATVAEF 32

RESULT 62
YIF9_YEAST STANDARD; PRT; 121 AA.
AC P40520;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Hypothetical 13.0 kDa protein in SNP1-GPPI intergenic region.
GN OrderedLocusNames=YII059C;
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RX MEDLINE=97313266; PubMed=9169870;
RA Churcher C.M., Bowman S., Badcock K., Bankier A., Brown D.,
RA Chillingworth T., Connor R., Devlin K., Gentles S., Hamlin N.,
RA Harris D.E., Horsnell T., Hunt S., Jagels K., Jones M., Lye G.,
RA Moule S., Odell C., Pearson D., Rajandream M.A., Rice P., Rowley N.,
RA Skelton J., Smith V., Walsh S., Whitehead S., Barrell B.G.;
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome IX.";
RL Nature 387:84-87(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C;
RA Marsischky G., Rolfs A., Richardson A., Kane M., Baqui M., Taycher E.,
RA Hu Y., Vannberg F., Weger J., Kramer J., Moreira D., Kelley P.,
RA Zuo D., Raphael J., Hogle C., Jepson D., Williamson J., Camargo A.,
RA Gonzaga L., Vasconcelos A.T., Simpson A., Kolodner R., Harlow E.,
RA LaBaer J.;
RT "Creation of the YFLEX clone resource: cloning of Saccharomycetes
cerevisiae ORFs in the Gateway recombinational cloning system.";
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
DR EMBL: Z38060; CAA86164.1; -.
DR EMBL: AY693258; AAT93277.1; -.
DR PIR: S48420; S48420.
DR Germline; 139596; -.
DR SGD: S00001321; YII059C.
KW Hypothetical protein; Transmembrane.
FT TRANSMEM 7 23 Potential.
SQ SEQUENCE 121 AA; 13030 MW; 18B23D652939CDD2 CRC64;

```

```
Query Match      81.8%; Score 18; DB 1; Length 121;
Best Local Similarity 57.1%; Pred. No. 2.2e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
DB 19 VTALAEF 25

RESULT 63
Q8TU11 ID Q8TU11 PRELIMINARY; PRT; 123 AA.
AC Q8TU11
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DE 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Predicted protein.
GN OrderedLocusNames=NA3803;
OS Methanosarcina acetivorans.
OC Archaea; Euryarchaeota; Methanomicrobia; Methanosarcinales;
OC Methanosarcinaceae; Methanosarcina.
OX NCBI_TaxID=2214;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C2A / ATCC 35395 / DSM 2834;
RX MEDLINE=21929760; PubMed=11932238; DOI=10.1101/gr.223902;
RA Galagan J.E., Nusbaum C., Roy A., Endrizzi M.G., Macdonald P.,
RA FitzHugh W., Calvo S., Engels R., Smirnov S., Atnoor D., Brown A.,
RA Allen N., Naylor J., Stange-Thomann N., Dearellano K., Johnson R.,
RA Linton L., McSwan P., McKernan K., Talamas J., Tirrell A., Ye W.,
RA Zimmer A., Barber R.D., Cann I., Graham D.E., Grahame D.A., Guss A.M.,
RA Hedderich R., Ingram-Smith C., Kuettnner H.C., Krzycki J.A.,
RA Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,
RA Springer T.A., Umayam L.A., White O., White R.H., de Macario E.C.,
RA Perry J.G., Jarrell K.F., Jing H., Macario A.J.L., Paulsen I.T.,
RA Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,
RA Metcalf W.W., Birren B.;
RT "The genome of Methanosarcina acetivorans reveals extensive metabolic
RT and physiological diversity.";
RL Genome Res. 12:532-542(2002).
DR ENBL; AE011092; ANM07154.1; -.
KW Complete proteome.
SQ SEQUENCE 123 AA; 12789 MW; D71653938D1DFB84 CRC64;

Query Match      81.8%; Score 18; DB 2; Length 123;
Best Local Similarity 57.1%; Pred. No. 2.2e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
DB 17 VEATAEF 23

RESULT 64
Q6MK11 ID Q6MK11 PRELIMINARY; PRT; 134 AA.
AC Q6MK11
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Chemotaxis response regulator CheY.
GN Names=cheY; OrderedLocusNames=Bd2409;
OS Bdellovibrio bacteriovorus.
OC Bacteria; Proteobacteria; Deltaproteobacteria; Bdellovibrionales;
OC Bdellovibrionaceae; Bdellovibrio.
OX NCBI_TaxID=959;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HD100 / DSM 50701 / ATCC 15356 / NCIB 9529;
RX PubMed=14752164; DOI=10.1126/science.1093027;
RA Rudulic S., Jagtap P., Rosinus A., Eppinger M., Baar C., Lanz C.,
RA Keller H., Lambert C., Evans K.J., Goemann A., Meyer F.,
RA Sockett R.E., Schuster S.C.;
```

```
RT "A predator unmasked: life cycle of Bdellovibrio bacteriovorus from a
RT genomic perspective.";
RL Science 303:689-692(2004).
DR ENBL; BX842652; CAE80226.1; -.
DR HSSP; P06143; IAB6.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0000156; F:two-component response regulator activity; IEA.
DR GO; GO:0000160; P:two-component signal transduction system (p. . .); IEA.
DR InterPro; IPR011006; CheY like.
DR InterPro; IPR001789; Response reg.
DR Pfam; PF00072; Response_reg; 1.
DR ProDom; PD000039; Response_reg; 1.
DR SMART; SM00448; REC; 1.
DR PROSITE; PS01110; RESPONSE REGULATORY; 1.
KW Complete proteome; Phosphorylation; Sensory transduction.
SQ SEQUENCE 134 AA; 15035 MW; A76AF72470EA58EF CRC64;

Query Match      81.8%; Score 18; DB 2; Length 134;
Best Local Similarity 57.1%; Pred. No. 2.4e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
DB 77 VRATAEF 83

RESULT 65
Q9NZ48 ID Q9NZ48 PRELIMINARY; PRT; 137 AA.
AC Q9NZ48
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Uncharacterized hematopoietic stem/progenitor cells protein
DE MDS026.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Hematopoietic stem/progenitor cells;
RA Zhao M., Gu J., Li N., Peng Y., Han Z., Chen Z.;
RA Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR ENBL; AF220046; AAF67639.1; -.
DR GO; GO:0000159; C:protein phosphatase type 2A complex; IEA.
DR GO; GO:0008601; F:protein phosphatase type 2A regulator activity; IEA.
DR GO; GO:0007165; P:signal transduction; IEA.
DR InterPro; IPR011044; Amine_DH_B_like.
DR InterPro; IPR000009; Pp2A_PR55.
DR PRINTS; PR00600; PP2APR55.
SQ SEQUENCE 137 AA; 14995 MW; 88B77FDD35F50F93 CRC64;

Query Match      81.8%; Score 18; DB 2; Length 137;
Best Local Similarity 57.1%; Pred. No. 2.4e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
DB 7 VITAEF 13

RESULT 66
Q6N7L1 ID Q6N7L1 PRELIMINARY; PRT; 141 AA.
AC Q6N7L1
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=KPA2246;
OS Rhodopsudomonas palustris.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
```

```

OC Bradyrhizobiaceae; Rhodopseudomonas.
OX NCBI_TaxID=1076;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CGA009 / ATCC BAA-98;
RX PubMed=14704707; DOI=10.1038/nbt923;
RA Larimer F.W., Chain P., Hauser L., Lamerdin J.E., Malfatti S., Do L.,
RA Land M.L., Pelletier D.A., Beatty J.T., Lang A.S., Tabita F.R.,
RA Gibson J.L., Hanson T.E., Bobst C., Torres y Torres J.L., Peres C.,
RA Harrison F.H., Gibson J., Harwood C.S.;
RT "Complete genome sequence of the metabolically versatile
RT photosynthetic bacterium Rhodopseudomonas palustris.";
RL Nat. Biotechnol. 22:55-61(2004).
DR EMBL; BX572600; CAE27687.1; -.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 141 AA; 15197 MW; 8E25F6A6E9C28C51 CRC64;

Query Match      81.8%; Score 18; DB 2; Length 141;
Best Local Similarity 57.1%; Pred. No. 2.5e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
Db 111 VSDSAEF 117

RESULT 67
Q41050
ID Q41050 PRELIMINARY; PRT; 146 AA.
AC Q41050;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Core protein.
OS Pisum sativum (Garden pea).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Viciaeae; Pisum.
OX NCBI_TaxID=3888;
RN [1]
RP SEQUENCE FROM N.A.
RC TSSUE=Leaf;
RX MEDLINE=97404425; PubMed=9256512; DOI=10.1073/pnas.94.17.9504;
RA Pohnmeyer K., Soil J., Steinkamp T., Hinman S., Wagner R.;
RT "Isolation and characterisation of an amino acid-selective channel
RT protein present in the chloroplastic outer envelope membrane.";
RL Proc. Natl. Acad. Sci. U.S.A. 94:9504-9509(1997).
DR EMBL; Z73553; CAA97910.1; -.
DR PIR; T06471; T06471.
DR GO; GO:0005743; C:mitochondrial inner membrane; IEA.
DR GO; GO:0008565; F:protein transporter activity; IEA.
DR GO; GO:0015031; P:protein transport; IEA.
DR InterPro; IPR003397; Tim17_Tim22.
DR Pfam; PF02466; Tim17; 1.
SQ SEQUENCE 146 AA; 15490 MW; B72E24ABF6A6C2AB CRC64;

Query Match      81.8%; Score 18; DB 2; Length 146;
Best Local Similarity 42.9%; Pred. No. 2.6e+03;
Matches 3; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
Db 135 IATAAEF 141

RESULT 68
Q9HNX0
ID Q9HNX0 PRELIMINARY; PRT; 147 AA.
AC Q9HNX0;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Vng1907h.

```

```

GN OrderedLocusNames=VNG1907H;
OS Halobacterium sp. (strain NRC-1 / ATCC 700922 / JCM 11081).
OC Archaea; Euryarchaeota; Halobacteriia; Halobacteriales;
OC Halobacteriaceae; Halobacterium.
OX NCBI_TaxID=64091;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20504483; PubMed=11016950; DOI=10.1073/pnas.190337797;
RA Ng W.V., Kennedy S.P., Mahairas G.G., Berquist B., Pan M.,
RA Shukla H.D., Lasky S.R., Baliga N.S., Thorsson V., Shrogha J.,
RA Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A.,
RA Leithausen B., Keller K., Cruz R., Danson M.J., Hough D.W.,
RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,
RA Isenbarger T.A., Peck R.P., Pohlischroder M., Spudich J.L., Jung K.-H.,
RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
RA Ebhardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.;
RT "Genome sequence of Halobacterium species NRC-1.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
DR EMBL; AS005089; AAG20100.1; -.
DR PIR; H84341; H84341.
KW Complete proteome.
SQ SEQUENCE 147 AA; 15759 MW; F83795F6D0D64204 CRC64;

Query Match      81.8%; Score 18; DB 2; Length 147;
Best Local Similarity 57.1%; Pred. No. 2.6e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
Db 102 VSSRAEF 108

RESULT 69
Q9CE23
ID Q9CE23 PRELIMINARY; PRT; 148 AA.
AC Q9CE23;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Non-heme iron-binding ferritin.
GN Name=dpsA; OrderedLocusNames=LL2025;
OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
OX NCBI_TaxID=1360;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IL1403;
RX MEDLINE=21235186; PubMed=11337471; DOI=10.1101/gr.1697R;
RA Bolotin A., Wincker P., Manger S., Jaillon O., Marme K.,
RA Weissenbach J., Ehrlich S.D., Sorokin A.;
RT "The complete genome sequence of the lactic acid bacterium Lactococcus
RT lactis ssp. lactis IL1403.";
RL Genome Res. 11:731-753(2001).
CC -!- SIMILARITY: Belongs to the dps family.
DR EMBL; AS006432; AAK08123.1; -.
DR PIR; A86878; A86878.
DR HSP; P80725; 1QGH.
DR GO; GO:0008199; F:ferric iron binding; IEA.
DR GO; GO:0006879; P:iron ion homeostasis; IEA.
DR GO; GO:0006950; P:response to stress; IEA.
DR InterPro; IPR002177; DPS.
DR InterPro; IPR009078; Ferritin/RR_like.
DR InterPro; IPR008331; Ferritin_Dps.
DR Pfam; PF0210; Ferritin; 1.
DR PRINTS; PR01346; HELNAPAPROT.
DR ProDom; PD149803; DPS; 1.
DR PROSITE; PS00818; DPS_1; 1.
KW Complete proteome.
SQ SEQUENCE 148 AA; 16653 MW; 8424544D9AFEB097 CRC64;

Query Match      81.8%; Score 18; DB 2; Length 148;
Best Local Similarity 57.1%; Pred. No. 2.6e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```
QY 1 VXXAAEF 7
|  |||
DB 69 VSTLAEF 75

RESULT 70
RL22_PICTO STANDARD; PRT; 149 AA.
AC Q6L1C2;
DT 25-OCT-2004 (Rel. 45, Created)
DT 25-OCT-2004 (Rel. 45, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE 50S ribosomal protein L22P.
GN Name=rpL22P; OrderedLocNames=PTO0645;
OS Picrophilus torridus.
OC Archaea; Euryarchaeota; Thermoplasmata; Thermoplasmatales;
OC Picrophilaceae; Picrophilus.
OX NCBI_TaxID=82076;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DSM 9790 / ATCC 700027;
RX PubMed=15184674; DOI=10.1073/pnas.0401356101;
RA Fueterer O., Angelov A., Liesegang H., Gottschalk G., Schleper C.,
RA Scheper B., Dock C., Antranikian G., Liebl W.;
RT "Genome sequence of Picrophilus torridus and its implications for life
around pH 0.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:9091-9096(2004).
CC -!- FUNCTION: This protein binds specifically to 23S rRNA. It makes
CC multiple contacts with different domains of the 23S rRNA in the
CC assembled 50S subunit and ribosome (By similarity).
CC -!- FUNCTION: The globular domain of the protein is located near the
CC polypeptide exit tunnel on the outside of the subunit, while an
CC extended beta-hairpin is found that lines the wall of the exit
CC tunnel in the center of the 70S ribosome. (By similarity).
CC -!- SUBUNIT: Part of the 50S ribosomal subunit.
CC -!- SIMILARITY: Belongs to the ribosomal protein L22p family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; A8017261; AAT43230.1; -.
DR HAMAP; MF01331; -.
DR PRODom; PD001032; Ribosomal L22; 1.
DR PROSITE; PS00464; RIBOSOMAL_L22; FALSE NEG.
KW Complete proteome; Ribosomal protein; RNA-binding; rRNA-binding.
SQ SEQUENCE 149 AA; 16731 MW; A80FA918151BA217 CRC64;

Query Match 81.8%; Score 18; DB 1; Length 149;
Best Local Similarity 57.1%; Pred. No. 2.6e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
|  |||
DB 92 VSAAEF 98

RESULT 71
Q89HB5 PRELIMINARY; PRT; 151 AA.
AC Q89HB5;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Blr6079 protein.
GN OrderedLocNames=blr6079;
OS Bradyrhizobium japonicum.
OC Bradyrhizobium japonicum.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
```

```
OC Bradyrhizobiaceae; Bradyrhizobium.
OX NCBI_TaxID=375;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=USDA110;
RX MEDLINE=22484998; PubMed=12597275;
RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiyama T.,
RA Sasamoto S., Watanabe A., Idesawa K., Iriguchi M., Kawashima K.,
RA Kohara M., Matsumoto M., Shimo S., Teurouka H., Wada T., Yamada M.,
RA Tabata S.;
RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
RT Bradyrhizobium japonicum USDA110.";
RL DNA Res. 9:189-197(2002).
DR EMBL; AP005957; BAC51344.1; -.
DR GO; GO:0045735; F:nutrient reservoir activity; IEA.
DR InterPro; IPR006045; Cupin.
DR InterPro; IPR007113; Cupin.
DR InterPro; IPR011051; RmlC_like_cupin.
DR Pfam; PF00190; Cupin; 1.
KW Complete proteome.
SQ SEQUENCE 151 AA; 16840 MW; 7AAAD1AC94717AC4 CRC64;

Query Match 81.8%; Score 18; DB 2; Length 151;
Best Local Similarity 57.1%; Pred. No. 2.6e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
|  |||
DB 134 VMAAEF 140

RESULT 72
FABZ CAUCR STANDARD; PRT; 159 AA.
AC Q9A7I4;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE (3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.-)
DE ((3R)-hydroxymyristoyl ACP dehydratase).
GN Name=fabZ; OrderedLocNames=CC1912;
OS Caulobacter crescentus.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacteriales;
OC Caulobacteraceae; Caulobacter.
OX NCBI_TaxID=155892;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 19089 / CB15;
RX MEDLINE=21173698; PubMed=11259647; DOI=10.1073/pnas.061029298;
RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
RA Eisen J.A., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
RA Potocka I., Nelson M.C., Newton A., Stephens C., Phadke N.D., Ely B.,
RA DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
RA Kolonay J.F., Smit J., Craven M.B., Khouri H.M., Shetty J.,
RA Berry K.J., Uitterback T.R., Tran K., Wolf A.M., Vamathevan J.J.,
RA Ermolaeva M.D., White O., Salzberg S.L., Venter J.C., Shapiro L.,
RA Fraser C.M.;
RT "Complete genome sequence of Caulobacter crescentus.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
CC -!- FUNCTION: Involved in saturated fatty acids biosynthesis.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the thioester dehydratase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; A8005865; AAK23887.1; -.
DR PIR; C87486; C87486.
```

```

DR TIGR; CC1912; -.
DR HAMAP; MF_00406; -. 1.
DR InterPro; IPR010084; FabZ.
DR InterPro; IPR006683; Thioestr_supf.
DR PFam; PF03061; 4HBT; 1.
DR TIGRFAMs; TIGR01750; fabZ; 1.
KW Complete proteome; Lipid A biosynthesis; Lipid synthesis; Lyase.
FT ACT_SITE 59 59 By similarity.
SQ SEQUENCE 159 AA; 17443 MW; 2C7EAA8668EA0141 CRC64;

Query Match      81.8%; Score 18; DB 1; Length 159;
Best Local Similarity 57.1%; Pred. No. 2.8e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
DB 142 VAAAEAF 148

RESULT 73
Q82RT3 PRELIMINARY; PRT; 171 AA.
AC Q82RT3;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein.
GN OrderedLocuNames=SAV60;
OS Streptomyces avermitilis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=33903;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MA-4680;
RX MEDLINE=22608306; PubMed=12692562;
RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T., Sakaki Y., Hattori M., Omura S.;
RT "Complete genome sequence and comparative analysis of the industrial microorganism Streptomyces avermitilis.";
RL Nat. Biotechnol. 21:526-531(2003).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=21477403; PubMed=11572948; DOI=10.1073/pnas.211433198;
RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C., Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osonoe T., Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
RT "Genome sequence of an industrial microorganism Streptomyces avermitilis: deducing the ability of producing secondary metabolites.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
DR EMBL; AP005021; BAC67769.1; -.
KW Complete proteome.
SQ SEQUENCE 171 AA; 18035 MW; 6CB7E15CDAC6FC75 CRC64;

Query Match      81.8%; Score 18; DB 2; Length 171;
Best Local Similarity 57.1%; Pred. No. 3e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
DB 43 VASIAEF 49

RESULT 74
Q9DDR2 PRELIMINARY; PRT; 173 AA.
AC Q9DDR2;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Protein phosphatase 2 regulatory subunit B (Fragment).

```

```

OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Intestine;
RA Minter R., Horton J.D., Watson M.D.;
RT "Development of Antibody Technology to Identify Natural Killer Cell Surface Antigens in Xenopus laevis.";
RL Thesis (1999), University of Durham, South Rd., Durham, UK.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Intestine;
RA Watson M.D.;
RA Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Contains 3 WD repeats.
DR EMBL; AF305084; AAG45003.1; -.
DR GO; GO:0000159; C:protein phosphatase type 2A complex; IEA.
DR GO; GO:0008601; F:protein phosphatase type 2A regulator activity; IEA.
DR GO; GO:0007165; P:signal transduction; IEA.
DR InterPro; IPR011044; Amine DH B like.
DR InterPro; IPR000009; Pp2A_Prs5.
DR InterPro; IPR001680; WD40.
DR Pfam; PF00400; WD40; 3.
DR PRINTS; PR00600; PP2APR55.
DR SMART; SM00320; WD40; 3.
DR PROSITE; PS01025; PR55_2; 1.
KW Repeat; WD repeat.
FT NON_TER 1
SQ SEQUENCE 173 AA; 20210 MW; BD84B23CA0F80880 CRC64;

Query Match      81.8%; Score 18; DB 2; Length 173;
Best Local Similarity 57.1%; Pred. No. 3e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
DB 64 VITAAEF 70

RESULT 75
Q8TQ97 PRELIMINARY; PRT; 179 AA.
AC Q8TQ97;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Intracellular protease.
GN Name=pfpi; OrderedLocuNames=MA1654;
OS Methanosarcina acetivorans.
OC Archaea; Euryarchaeota; Methanomicrobia; Methanosarcinales;
OC Methanosarcinaceae; Methanosarcina.
OX NCBI_TaxID=2214;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C2A / ATCC 35395 / DSM 2834;
RX MEDLINE=21929760; PubMed=11932238; DOI=10.1101/gr.223902;
RA Galagan J.E., Nusbaum C., Roy A., Smirnov S., Atnoor D., Brown A., FitzHugh W., Calvo S., Engels R., Endrizzi M.G., Macdonald P., Allen N., Naylor J., Stange-Thomann N., DeArelano K., Johnson R., Linton L., McEwan P., McKernan K., Talamas J., Firrell A., Ye W., Zimmer A., Barber R.D., Cann I., Graham D.E., Grahame D.A., Guss A.M., Hedderich R., Ingram-Smith C., Kuettner H.C., Krzycki J.A., Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K., Perry J.G., Jarrell K.F., White O., White R.H., de Macario E.C., Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E., Metcalf W.W., Birren B.;
RT "The genome of Methanosarcina acetivorans reveals extensive metabolic and physiological diversity.";
RL Genome Res. 12:532-542(2002).

```



```

DR EMBL; AE010838; AAM05062.1; -.
DR HSSP; O59413; 1G21.
DR GO; GO:0008233; P:peptidase activity; IEA.
DR InterPro; IPR002818; ThiJ/Pfpl.
DR Pfam; PF01965; DJ-1_Pfpl; 1.
KW Complete proteome; Protease.
SQ SEQUENCE 179 AA; 19359 MW; E0FDC30D4BF4F456 CRC64;

Query Match      81.8%; Score 18; DB 2; Length 179;
Best Local Similarity 57.1%; Pred. No. 3.1e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VXXXXAEF 7
Db 38 VTVAEEF 44

RESULT 76
Q6AC19 Q6AC19 PRELIMINARY; PRT; 184 AA.
AC Q6AC19;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=Lxx24480;
OS Leifsonia xyli (subsp. xyli).
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Micrococciaceae; Microbacteriaceae; Leifsonia.
OX NCBI_TaxID=59736;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CTCB07;
RX PubMed=15305603;
RA Monteiro-Vitorello C.B., Camargo L.E.A., Van Sluys M.A.,
RA Kitajima J.P., Truffi D., do Amaral A.M., Harakava R.,
RA de Oliveira J.C.F., Wood D., de Oliveira M.C., Miyaki C.Y.,
RA Takita M.A., da Silva A.C.R., Furlan L.R., Carraro D.M., Camarotte G.,
RA Almeida N.F. Jr., Carier H., Coutinho L.L., El-Dorry H.A.,
RA Ferro M.I.T., Gagliardi P.R., Gigliotti E., Goldman M.H.S.,
RA Goldman G.H., Kimura E.T., Ferro E.S., Kuramae E.E., Lemos E.G.M.,
RA Lemos M.V.F., Mauro S.M.Z., Machado M.A., Marino C.L., Menck C.F.,
RA Nunes L.R., Oliveira R.C., Pereira G.G., Siqueira W., de Souza A.A.,
RA Tsai S.M., Zanca A.S., Simpson A.J.G., Brumbley S.M., Setubal J.C.;
RT "The genome sequence of the Gram-positive sugarcane pathogen Leifsonia
RT xyli subsp. xyli."
RL Mol. Plant Microbe Interact. 17:827-836(2004).
DR EMBL; AE016822; AAT90073.1; -.
KW Complete proteome.
SQ SEQUENCE 184 AA; 18354 MW; 21557D7FF637DB99 CRC64;

Query Match      81.8%; Score 18; DB 2; Length 184;
Best Local Similarity 42.9%; Pred. No. 3.2e+03;
Matches 3; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VXXXXAEF 7
Db 89 IAATAEF 95

RESULT 77
Q76YY3 Q76YY3 PRELIMINARY; PRT; 186 AA.
AC Q76YY3;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
GN ORFNames=Aeh1ORF102C;
OS Bacteriophage Aeh1.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae;
OC T4-like viruses.
OX NCBI_TaxID=227470;

RN [1]
RP SEQUENCE FROM N.A.
RA Petrov V., Nolan J., Bertrand C., Letarov A.V., Krich H.M.,
RA Karam J.D.;
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Petrov V., Nolan J., Karam J.D.;
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY266303; AAQ17763.1; -.
KW Hypothetical protein.
SQ SEQUENCE 186 AA; 21839 MW; 45060CC5F8B03B11 CRC64;

Query Match      81.8%; Score 18; DB 2; Length 186;
Best Local Similarity 57.1%; Pred. No. 3.2e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VXXXXAEF 7
Db 126 WVATAEF 132

RESULT 78
Q9N416 Q9N416 PRELIMINARY; PRT; 189 AA.
AC Q9N416;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein Y71F9AL.10.
GN Name=Y71F9AL.10; ORFNames=Y71F9AL.10;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RG WormBase Consortium;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium."
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Bradshaw-Cordum H., Scott K., Graves T.;
RT "The sequence of C. elegans cosmid Y71F9AL."
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Wilson R.;
RL Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Wilson R.;
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RG WormBase Consortium;
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Contains 1 RING-type zinc finger.
DR EMBL; AC024200; AAF35997.2; -.
DR HSSP; Q9LRB7; 1IYM.
DR WormBase; WBGene00022115; Y71F9AL.10.

```

```

DR WormRep; Y71F9AL.10; CE31386.
DR GO; GO:0000151; C:ubiquitin ligase complex; IEA.
DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0016567; P:protein ubiquitination; IEA.
DR InterPro; IPR001841; Znf ring.
DR Pfam; PF00097; zf-C3HC4; 1.
DR SMART; SM00184; RING; 1.
DR PROSITE; PSS00089; ZF_RING_2; 1.
KW Hypothetical protein; Metal-binding; Zinc; Zinc-finger.
SQ SEQUENCE 189 AA; 21048 MW; 7BDA2D05F0362CD0 CRC64;

Query Match 81.8%; Score 18; DB 2; Length 189;
Best Local Similarity 57.1%; Pred. No. 3.2e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
DB 37 VMTAAEF 43

RESULT 79
Q6CDT8 PRELIMINARY; PRT; 193 AA.
AC Q6CDT8
DT 23-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Similarity.
GN ORFNames=YALI0B212749;
OS Yarrowia lipolytica CLIB99.
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Dipodascaceae; Yarrowia.
OX NCBI_TaxID=284591;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CLIB99;
RG Genolevures;
RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,
RA Lafontaine I., de Montigny J., Marck C., Neuveglise C., Talla E.,
RA Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V.,
RA Barnay S., Blanchin S., Beckerich J.M., Beyne E., Bleykasten C.,
RA Boisrame A., Boyer J., Cattolico L., Confanioleri F., de Daruvar A.,
RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,
RA Hentraye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,
RA Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,
RA Nicaud J.M., Nikolski M., Oztas S., Ozier-Kalogeropoulos O.,
RA Pellenz S., Potier S., Richard G.F., Straub M.L., Suleau A.,
RA Swennene D., Tekai F., Wesolowski-Louvel M., Westhof E., Wirth B.,
RA Zeniou-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A.,
RA Bouchier C., Caudron B., Scarpelli C., Gallardin C., Weissenbach J.,
RA Wincker P., Souciet J.L.;
RT "Genome evolution in yeasts.";
RN Nature 430:35-44(2004).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CLIB99;
RA Genoscope;
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CR382128; CAG83427.1; -.
SQ SEQUENCE 193 AA; 22016 MW; 939292835977B6BF CRC64;

Query Match 81.8%; Score 18; DB 2; Length 193;
Best Local Similarity 57.1%; Pred. No. 3.3e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
DB 86 VATLAEF 92

RESULT 80
Q8E1V1 PRELIMINARY; PRT; 193 AA.
AC Q8E1V1;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Membrane protein, putative.
DE OrderedLocusNames=SAG0250;
GN Streptococcus agalactiae (serotype V).
OS Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=216466;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=2603 V/R / Serotype V;
RX MEDLINE=22222988; PubMed=12200547; DOI=10.1073/pnas.182380799;
RA Tettelin H., Masiugnani V., Cieslewicz M.J., Eisen J.A., Peterson S.N.,
RA Wessels M.R., Paulsen I.T., Nelson K.E., Margarit I., Read T.D.,
RA Madoff L.C., Wolf A.M., Beanan M.J., Brinkac L.M., Daugherty S.C.,
RA DeBoy R.T., Durkin A.S., Kolonay J.F., Madupu R., Lewis M.R.,
RA Radune D., Fedorova N.B., Scanlan D., Khouri H.M., Mulligan S.,
RA Carty H.A., Cline R.T., Van Aken S.E., Gill J., Scarselli M., Mora M.,
RA Iacobini E.T., Brettoni C., Galli G., Mariani M., Vegni F., Malone D.,
RA Rinaldo D., Rappuoli R., Telford J.L., Kasper D.L., Grandi G.,
RA Fraser C.M.;
RT "Complete genome sequence and comparative genomic analysis of an
RT emerging human pathogen, serotype V Streptococcus agalactiae.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:12391-12396(2002).
DR EMBL; AR014203; AAM99157.1; -.
DR TIGR; SAG0250; -.
KW Complete proteome.
SQ SEQUENCE 193 AA; 22090 MW; 06D80D9E108F8159 CRC64;

Query Match 81.8%; Score 18; DB 2; Length 193;
Best Local Similarity 57.1%; Pred. No. 3.3e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
DB 41 VTSIAEF 47

RESULT 81
Q63EZ7 PRELIMINARY; PRT; 194 AA.
AC Q63EZ7;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Group-specific protein.
GN ORFNames=B7ZK0912;
OS Bacillus cereus ZK.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=288681;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ZK;
RA Brettin T.S., Bruce D., Challacombe J.F., Gilna P., Han C., Hill K.,
RA Hitchcock P., Jackson P., Keim P., Longmire J., Lucas S., Okinaka R.,
RA Richardson P., Rubin E., Tice H.;
RT "Complete genome sequence of Bacillus cereus ZK.";
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CP000001; AAU19332.1; -.
SQ SEQUENCE 194 AA; 22360 MW; CCCBCE4F1C508863 CRC64;

Query Match 81.8%; Score 18; DB 2; Length 194;
Best Local Similarity 57.1%; Pred. No. 3.3e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
DB 43 VOAAAEF 49

```

```

RESULT 82
O54133 PRELIMINARY; PRT; 194 AA.
AC O54133;
DT 01-JUN-1998 (TREMBlrel. 06, Created)
DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
DE 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Hypothetical protein SC05867.
GN ORFNames=SC2E9.08;
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RX MEDLINE=21996410; PubMed=12000953; DOI=10.1038/417141a;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Lark L., Murphy L.D., Oliver K., O'Neil S.,
RA Rabinowitsch E., Rajandream M.A., Rutherford K.M., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J.R., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2)." / 147(2002).
RL Nature 417:141-147(2002).
DR EMBL; AL939125; CAAL16476.1; -.
DR PIR; T34819; T34819.
DR GO; GO:0003824; F: catalytic activity; IEA.
DR InterPro; IPR006683; Thioestr_supf.
DR Pfam; PF03061; 4HBT; 1.
DR Complete proteome; Hypothetical protein.
KW SEQUENCE 194 AA; 20474 MW; 9FE391B9E3CED80E CRC64;
SQ SEQUENCE FROM N.A.

Query Match 81.8%; Score 18; DB 2; Length 194;
Best Local Similarity 57.1%; Pred. No. 3.3e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 VXXAAEF 7
DB 58 VSITAEF 64

RESULT 83
O82KK3 PRELIMINARY; PRT; 194 AA.
AC O82KK3;
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DE 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=SAV2400;
OS Streptomyces avermitilis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=33903;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MA-4680;
RX MEDLINE=21477403; PubMed=11572948; DOI=10.1073/pnas.211433198;
RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA Shinose H., Takahashi Y., Horikawa H., Nakazawa H., Osone T.,
RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
RT "Genome sequence of an industrial microorganism Streptomyces
RT avermitilis: deducing the ability of producing secondary
RT metabolites.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MA-4680;

```

```

RX MEDLINE=22608306; PubMed=12692562;
RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
RA Sakaki Y., Hattori M., Omura S.;
RT "Complete genome sequence and comparative analysis of the industrial
RT microorganism Streptomyces avermitilis.";
RL Nat. Biotechnol. 21:526-531(2003).
DR EMBL; AF005030; BAC70111.1; -.
DR GO; GO:0003824; F: catalytic activity; IEA.
DR InterPro; IPR006683; Thioestr_supf.
DR Pfam; PF03061; 4HBT; 1.
DR Complete proteome; Hypothetical protein.
KW SEQUENCE 194 AA; 20316 MW; 3BE41F14E45228D5 CRC64;
SQ SEQUENCE FROM N.A.

Query Match 81.8%; Score 18; DB 2; Length 194;
Best Local Similarity 57.1%; Pred. No. 3.3e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 VXXAAEF 7
DB 58 VSITAEF 64

RESULT 84
O7NNX1 PRELIMINARY; PRT; 198 AA.
AC O7NNX1;
DT 01-MAR-2004 (TREMBlrel. 26, Created)
DT 01-MAR-2004 (TREMBlrel. 26, Last sequence update)
DE 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Molybdopterin-guanine dinucleotide biosynthesis protein.
GN Names=mba; OrderedLocusNames=CV1625;
OS Chromobacterium violaceum.
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Chromobacterium.
OX NCBI_TaxID=536;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 12472 / DSM 30191;
RX MEDLINE=22892880; PubMed=14500782; DOI=10.1073/pnas.1832124100;
RA Vasconcelos A.T.R., de Almeida D.F., Hungria M., Guimaraes C.T.,
RA Antonio R.V., Almeida F.C., de Almeida L.G.P., de Almeida R.,
RA Alves-Gomes J.A., Andrade E.M., Araripe J., de Araujo M.F.F.,
RA Astolfi-Filho S., Azevedo V., Baptista A.J., Batus L.A.M.,
RA Batista J.S., Belo A., van den Berg C., Bogo M., Bonatto S.,
RA Bordignon J., Brígido M.M., Brito C.A., Brocchi M., Burity H.A.,
RA Camargo A.A., Cardoso D.D.P., Carneiro N.P., Carraro D.M.,
RA Carvalho C.M.B., Cascardo J.C.M., Cavada B.S., Chueire L.M.O.,
RA Creczynski-Pasa T.B., Cunha-Junior N.C., Fagundes N., Falcao C.L.,
RA Fantinatti F., Farias I.P., Felipe M.S.S., Ferrari L.P., Ferro J.A.,
RA Ferro M.I.T., Franco G.R., Freitas N.S.A., Furlan L.R.,
RA Gazzinelli R.T., Gomes E.A., Goncalves P.R., Grangeiro T.B.,
RA Grattapaglia D., Grisard E.C., Hanna E.S., Jardim S.N., Laurino J.,
RA Leao L.C.T., Lima L.F.A., Loureiro M.F., Lyra M.C.P.P.,
RA Madeira H.M.F., Manfio G.P., Maranhao A.Q., Martins W.S.,
RA di Mauro S.M.Z., de Medeiros S.R.B., Meissner R.V., Moreira M.A.M.,
RA Nascimento P.F., Nicolas M.F., Oliveira J.G., Oliveira S.C.,
RA Paixao R.F.C., Parente J.A., Pedrosa F.O., Pena S.D.J., Pereira J.O.,
RA Pereira M., Pinto L.S.R.C., Pinto L.S., Porto J.I.R., Potrich D.P.,
RA Ramalho-Neto C.E., Reis A.M.M., Rigo L.U., Rondinelli E.,
RA Santos E.B.P., Santos F.R., Schneider M.P.C., Seuneh H.N.,
RA Silva A.M.R., da Silva A.L.C., Silva D.W., Silva R., Simoes I.C.,
RA Simon D., Soares C.M.A., Soares R.B.A., Souza E.M., Souza K.R.L.,
RA Souza R.C., Steffens M.B.R., Steindel M., Teixeira S.R., Urmenyi T.,
RA Vettore A., Wassem R., Zaha A., Simpson A.J.G.;
RT "The complete genome sequence of Chromobacterium violaceum reveals
RT remarkable and exploitable bacterial adaptability.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:11660-11665(2003).
DR EMBL; AE016915; AAQ59301.1; -.
DR HSSP; P32173; 1E5K.
DR Complete proteome.
KW SEQUENCE 198 AA; 20709 MW; B1670D8D6C782B0C CRC64;
SQ SEQUENCE FROM N.A.

Query Match 81.8%; Score 18; DB 2; Length 198;

```

Best Local Similarity 57.1%; Pred. No. 3.4e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
|
71 VTDAAEF 77

RESULT 85

ID RS6 TOBAC STANDARD; PRT; 199 AA.
AC P29345;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE 40S ribosomal protein S6 (Fragment).
GN Name:RPS6;
OS Nicotiana tabacum (Common tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC lamiales; Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=4097;
RN [1]
RC SEQUENCE FROM N.A.
RP STRAIN=cv. SRL; TISSUE=Leaf;
RX MEDLINE=93027271; PubMed=1408841;
RA Hansen G., Estruch J.J., Spena A.;
RT "Tobacco cDNA encoding the ribosomal protein S6.";
RL Nucleic Acids Res. 20:5230-5230(1992).
CC -I- FUNCTION: May play an important role in controlling cell growth
CC and proliferation through the selective translation of particular
CC classes of mRNA.
CC -I- PTM: Ribosomal protein S6 is the major substrate of protein
CC kinases in eukaryote ribosomes.
CC -I- SIMILARITY: Belongs to the ribosomal protein S6e family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; X68050; CAA48187.1; ALT INIT.
DR InterPro; IPR001377; Ribosomal_S6e.
DR Pfam; PF01092; Ribosomal_S6e; 1.
DR ProDom; PD003460; Ribosomal_S6e; 1.
DR PROSITE; PS00578; RIBOSOMAL_S6E; 1.
KW Phosphorylation; Ribosomal protein.
FT NON_TER 1 1
SQ SEQUENCE 199 AA; 22620 MW; 03EEBA0BD87B77F5 CRC64;

Query Match 81.8%; Score 18; DB 1; Length 199;
Best Local Similarity 57.1%; Pred. No. 3.4e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
|
112 VNTTAEF 118

RESULT 86

Q9NBA9 PRELIMINARY; PRT; 201 AA.
ID Q9NBA9
AC Q9NBA9
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Stretchin-MLCK (Fragment).
GN Name=Strn-MLck;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

Query Match 81.8%; Score 18; DB 2; Length 199;
Best Local Similarity 57.1%; Pred. No. 3.4e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
|
112 VNTTAEF 118

OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RC SEQUENCE FROM N.A.
RP MEDLINE=20351481; PubMed=10891286; DOI=10.1006/jmbi.2000.3802;
RX Champagne M.B., Edwards K.A., Erickson H.P., Kiehart D.P.;
RA "Drosophila stretchin-MLCK is a novel member of the Titin/Myosin light
RT chain kinase family.";
RL J. Mol. Biol. 300:759-777(2000).
DR EMBL; AF257311; AAF90129.1; --
DR HSSP; Q3UOH9; 1DJS.
DR FlyBase; FBgn0013988; Strn-MLck.
DR GO; GO:0004685; F:calcium- and calmodulin-dependent protein k. . .; IDA.
DR GO; GO:0004687; F:myosin-light-chain kinase activity; IDA.
DR InterPro; IPR003599; IG.
DR InterPro; IPR007110; IG-like.
DR SMART; SM00409; IG; 2.
DR PROSITE; PS50835; IG_LIKE; 1.
FT NON_TER 1 1
FT NON_TER 201 201
SQ SEQUENCE 201 AA; 22150 MW; CFC118FDC3C423DD CRC64;

Query Match 81.8%; Score 18; DB 2; Length 201;
Best Local Similarity 57.1%; Pred. No. 3.4e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
|
14 VAISAEF 20

RESULT 87

Q48517 PRELIMINARY; PRT; 202 AA.
AC Q48517;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Ddl protein (Fragment).
GN Name=ddl;
OS Lactobacillus confusus.
OC Bacteria; Firmicutes; Lactobacillales; Weissella.
OX NCBI_TaxID=1583;
RN [1]
RC SEQUENCE FROM N.A.
RP STRAIN=CIP102578;
RX MEDLINE=95129920; PubMed=7828933; DOI=10.1016/0378-1119(94)00692-L;
RA Elisha B.G., Courvalin P.;
RT "Analysis of genes encoding D-alanine:D-alanine ligase-related enzymes
RL in Leuconostoc mesenteroides and Lactobacillus spp.";
RL Gene 152:79-83(1995).
DR EMBL; U08910; AAA87680.1; --
DR HSSP; P71454; 1EHI.
DR InterPro; IPR011095; Dala_Dala_lig_C.
DR InterPro; IPR011127; Dala_Dala_lig_N.
DR Pfam; PF07478; Dala_Dala_lig_C; 1.
DR Pfam; PF01820; Dala_Dala_lig_N; 1.
FT NON_TER 1 1
FT NON_TER 202 202
SQ SEQUENCE 202 AA; 22335 MW; 702BB09B67FF1A49 CRC64;

Query Match 81.8%; Score 18; DB 2; Length 202;
Best Local Similarity 57.1%; Pred. No. 3.4e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
|
81 VTNAAEF 87

RESULT 88

Q6SKE7

```

ID Q6SKE7 PRELIMINARY; PRT; 207 AA.
AC Q6SKE7;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
OS Arthrobacter aureus.
OS Plasmid pAA1.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Micrococineae; Micrococcaceae; Arthrobacter.
OX NCBI_TaxID=43663;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TC1;
RX PubMed=15240330; DOI=10.1128/AEM.70.7.4402-4407.2004;
RA Sajjaphan K., Shapir N., Wackett L.P., Palmer M., Blackmon B.,
RA Tomkins J., Sadowsky M.J.;
RT "Arthrobacter aureusens TC1 Atrazine Catabolism Genes trzN, atzB, and
RT atzC Are Linked on a 160-Kilobase Region and Are Functional in
RT Escherichia coli.";
RL Appl. Environ. Microbiol. 70:4402-4407(2004).
DR EMBL: AY456696; AAS20024.1; -.
DR InterPro: IPR001763; Rhodanese-like.
DR Pfam: PF00581; Rhodanese; 1.
DR SMART: SM00450; RHOD; 1.
DR PROSITE: PSS0206; RHODANES_3; 1.
KW Hypothetical protein; Plasmid.
SQ SEQUENCE 207 AA; 21741 MW; F10397D13AD2F10A CRC64;

Query Match 81.8%; Score 18; DB 2; Length 207;
Best Local Similarity 57.1%; Pred. No. 3.5e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
Db 33 VRSAAEF 39

RESULT 89
Q64XJ4 PRELIMINARY; PRT; 211 AA.
ID Q64XJ4
AC Q64XJ4;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein.
GN ORFNames=BP1032;
OS Bacteroides fragilis.
OC Bacteria; Bacteroidetes; Bacteroides (class); Bacteroidales;
OC Bacteroidaceae; Bacteroides.
OX NCBI_TaxID=817;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=YCH46;
RA Kuwahara T., Yamashita A., Hirakawa H., Nakayama H., Toh H., Okada N.,
RA Kuwaha S., Hattori M., Hayashi T., Ohnishi Y.;
RT "Genomic analysis of Bacteroides fragilis reveals extensive DNA
RT inversions regulating cell surface adaptation.";
RL Proc. Natl. Acad. Sci. U.S.A. 0:0-0(2004).
DR EMBL: AP006841; BAD47782.1; -.
KW Hypothetical protein.
SQ SEQUENCE 211 AA; 24442 MW; C86EC7A0D93BA917 CRC64;

Query Match 81.8%; Score 18; DB 2; Length 211;
Best Local Similarity 57.1%; Pred. No. 3.6e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
Db 17 VTVAEEF 23

RESULT 90

```

```

Q6VTX4 PRELIMINARY; PRT; 211 AA.
ID Q6VTX4
AC Q6VTX4;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
OS Choriostoeura fumiferana defective nucleopolyhedrovirus.
OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
OC Nucleopolyhedrovirus.
OX NCBI_TaxID=74660;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20149221; PubMed=10683327; DOI=10.1006/viro.1999.0138;
RA Li X., Barrett J., Pang A., Klose R.J., Krell P.J., Arif B.M.;
RT "Characterization of an overexpressed spindle protein during a
RT baculovirus infection.";
RL Virology 268:56-67(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Lauzon H.A.M., Jamieson P.B., Krell P.J., Arif B.M.;
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY327402; AAQ91712.1; -.
DR InterPro: IPR009661; DUF1251.
DR Pfam: PF06856; DUF1251; 1.
KW Hypothetical protein.
SQ SEQUENCE 211 AA; 23673 MW; A24B5286DF93CA3A CRC64;

Query Match 81.8%; Score 18; DB 2; Length 211;
Best Local Similarity 57.1%; Pred. No. 3.6e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
Db 30 VSADAEF 36

RESULT 91
Q6CE44 PRELIMINARY; PRT; 212 AA.
ID Q6CE44
AC Q6CE44;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Similarity.
GN CFNames=YALI0B186789;
OS Yarrowia lipolytica CLIB99.
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Dipodascaceae; Yarrowia.
OX NCBI_TaxID=284591;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CLIB99;
RG Genolevures;
RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,
RA Lafontaine I., de Montigny J., Marck C., Neugeglise C., Talla E.,
RA Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V.,
RA Barnay S., Blanchin S., Beckerich J.M., Beyne E., Bleykasten C.,
RA Boisrame A., Boyer J., Cattolico L., Confanioli F., de Daruvar A.,
RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,
RA Hantraye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,
RA Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,
RA Nicaud J.M., Nikolski M., Ostas S., Ozier-Kalogeropoulos O.,
RA Pellenz S., Potier S., Richard G.F., Straub M.L., Suleau A.,
RA Zennaro D., Tekala F., Wesolowski-Louvel M., Westhof E., Wirth B.,
RA Zeniou-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A.,
RA Bouchier C., Caudron B., Scarpelli C., Gaillardin C., Weissenbach J.,
RA Wincker P., Souciet J.L.;
RT "Genome evolution in yeasts.";
RL Nature 430:35-44(2004).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CLIB99;

```

```
RA Genoscope;
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CR382128; CAG83321.1; -.
SQ SEQUENCE 212 AA; 24023 MW; 95219C003E2B41E2 CRC64;

Query Match
Best Local Similarity 81.8%; Score 18; DB 2; Length 212;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
| | | |
Db 180 VLAAEF 186

RESULT 92
Q9NBBO PRELIMINARY; PRT; 212 AA.
ID Q9NBBO;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Stretchin-MLCK (Fragment).
GN Name=Strn-Mlck;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20351481; PubMed=10891286; DOI=10.1006/jmbi.2000.3802;
RA Champagne M.B., Edwards K.A., Erickson H.P., Kienhart D.P.;
RT "Drosophila stretchin-MLCK is a novel member of the Titin/Myosin light
chain kinase family.";
RL J. Mol. Biol. 300:759-777(2000).
DR EMBL; AF257310; AAF90128.1; -.
DR HSP; Q9UQH9; IDJ5
DR FlyBase; FBgn0013988; Strn-Mlck.
DR GO; GO:0004685; F:calcium- and calmodulin-dependent protein k. . ; IDA.
DR GO; GO:0004687; F:myosin-light-chain kinase activity; IDA.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR007110; Ig-like.
DR SMART; SM00409; IG; 2
DR PROSITE; PS50835; IG-LIKE; 1.
FT NON_TER 1
FT NON_TER 212
SQ SEQUENCE 212 AA; 23599 MW; 5DC06D88FD2B4803 CRC64;

Query Match
Best Local Similarity 81.8%; Score 18; DB 2; Length 212;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
| | | |
Db 25 VAISAEF 31

RESULT 93
Q9PPH4 PRELIMINARY; PRT; 212 AA.
ID Q9PPH4;
AC Q9PPH4;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein Cj0733.
GN OrderedLocusNames=Cj0733;
OS Campylobacter jejuni.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Campylobacteraceae; Campylobacter.
OX NCBI_TaxID=197;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=NCTC 11168;
```

```
RX MEDLINE=20150912; PubMed=10688204; DOI=10.1038/35001088;
RA Parkhill J., Wren B.W., Mungall K.L., Ketley J.M., Churcher C.M.,
RA Basham D., Chillingworth T., Davies R.M., Feltwell T., Holtroyd S.,
RA Jagels K., Karlyshev A.V., Moule S., Pallen M.J., Penn C.W.,
RA Quail M.A., Rajandream M.A., Rutherford K.M., van Vliet A.H.M.,
RA Whitehead S., Barrell B.G.;
RT "The genome sequence of the food-borne pathogen Campylobacter jejuni
reveals hypervariable sequences.";
RL Nature 403:665-668(2000).
DR EMBL; AL139076; CAB73007.1; -.
DR PIR; H81344; H81344.
DR GO; GO:0003824; P:catalytic activity; IEA.
DR GO; GO:0016791; P:phosphoric monoester hydrolase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR005834; Dehal-like hydro.
DR InterPro; IPR006383; HAD_SF_IB.
DR InterPro; IPR006385; HAD_SF_IB_hyp2.
DR Pfam; PF00702; Hydrolase_1.
DR TIGRFAMs; TIGR01488; HAD-SF-IB; 1.
DR TIGRFAMs; TIGR01490; HAD-SF-IB-hypl; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 212 AA; 25015 MW; 8A892FEEBDF54ED4 CRC64;

Query Match
Best Local Similarity 81.8%; Score 18; DB 2; Length 212;
Matches 3; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
| | | |
Db 105 ISATAEF 111

RESULT 94
Q97GY8 PRELIMINARY; PRT; 213 AA.
ID Q97GY8;
AC Q97GY8;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Phosphoserine phosphatase family enzyme.
GN OrderedLocusNames=CAC2227;
OS Clostridium acetobutylicum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1488;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
RX MEDLINE=21359325; PubMed=11466286;
DOI=10.1128/JB.183.16.4823-4838.2001;
RA Noelling J., Breton G., Omeichenko M.V., Makarova K.S., Zeng Q.,
RA Gibson R., Lee H.M., Dubois J., Qiu D., Hitti J., Wolf Y.I.,
RA Tatusov R.L., Sabathe F., Doucette-Stamm L.A., Soucaille P.,
RA Daly M.J., Bennett G.N., Koonin E.V., Smith D.R.;
RT "Genome sequence and comparative analysis of the solvent-producing
bacterium Clostridium acetobutylicum.";
RL J. Bacteriol. 183:4823-4838(2001).
DR EMBL; AE007723; AAK80184.1; -.
DR PIR; E97174; E97174.
DR GO; GO:0016791; P:phosphoric monoester hydrolase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR006383; HAD_SF_IB.
DR InterPro; IPR006385; HAD_SF_IB_hyp2.
DR TIGRFAMs; TIGR01488; HAD-SF-IB; 1.
DR TIGRFAMs; TIGR01490; HAD-SF-IB-hypl; 1.
KW Complete proteome.
SQ SEQUENCE 213 AA; 25052 MW; 16F92EA612343A40 CRC64;

Query Match
Best Local Similarity 81.8%; Score 18; DB 2; Length 213;
Matches 3; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
```

```

KW Complete proteome.
SQ SEQUENCE 216 AA; 22587 MW; 423CD8B32A67B7A8 CRC64;

Query Match 81.8%; Score 18; DB 2; Length 216;
Best Local Similarity 57.1%; Pred. No. 3.7e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
DB 70 VAAAEF 76

RESULT 97
Q7PCW7 PRELIMINARY; PRT; 217 AA.
AC Q7PCW7;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DE 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE AGCP6282 (ENSANGP0000018110).
GN Name=agc52667; Synonyms=ENSANGG0000015621;
GN ORFNames=ENSANGG0000015279;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.
OX NCBI_TaxID=180454;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PEST;
RA Anopheles Genome Sequencing Consortium;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=PEST;
RA Anopheles Genome Sequencing Consortium;
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAB01008960; EAA11056.1; -.
SQ SEQUENCE 217 AA; 24420 MW; F5A1CC770B728ECB CRC64;

Query Match 81.8%; Score 18; DB 2; Length 217;
Best Local Similarity 57.1%; Pred. No. 3.7e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
DB 59 VAAAEF 65

RESULT 98
Q7MTV9 PRELIMINARY; PRT; 217 AA.
AC Q7MTV9;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DE 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=PGI823;
OS Porphyromonas gingivalis (Bacteroides gingivalis).
OC Bacteria; Bacteroidetes; Bacteroides (class); Bacteroidales;
OC Porphyromonadaceae; Porphyromonas.
OX NCBI_TaxID=837;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=W83;
RX MEDLINE=22829867; PubMed=12949112;
RX DOI=10.1128/JB.185.18.5591-5601.2003;
RA Nelson K.E., Fleischmann R.D., DeBoy R.T., Paulsen I.T., Fouts D.E.,
RA Eisen J.A., Dougherty S.C., Dodson R.J., Durkin A.S., Gwinn M.L.,
RA Haft D.H., Kolonay J.F., Nelson W.C., Mason T.M., Tallon L., Gray J.,
RA Granger D., Tettelin H., Dong H., Galvin J.L., Duncan M.J.,

DB 110 ISAAEF 116

RESULT 95
Q9N630 PRELIMINARY; PRT; 215 AA.
AC Q9N630;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Stretchin-MLCK (Fragment).
GN Name=Strn-Mlck;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20351481; PubMed=10891286; DOI=10.1006/jmbi.2000.3802;
RA Champagne M.B., Edwards K.A., Erickson H.P., Kiehart D.P.;
RT "Drosophila stretchin-MLCK is a novel member of the Titin/Myosin light
RT chain kinase family.";
RL J. Mol. Biol. 300:759-777(2000).
DR EMBL; AF257312; AAF90130.1; -.
DR EMBL; AF254365; AAF90114.1; -.
DR FlyBase; FBgn0013988; Strn-Mlck.
DR GO; GO:0004685; F-actin- and calmodulin-dependent protein k. . . ; IDA.
DR GO; GO:0004687; F-actin- and calmodulin-dependent protein k. . . ; IDA.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003598; Ig_c2.
DR SMART; SM00408; IGC2; 1.
DR PROSITE; PS50835; IG_LIKE; 2.
FT NON_TER 215
FT NON_TER 215
SQ SEQUENCE 215 AA; 24254 MW; E6A8019E1983E9FC CRC64;

Query Match 81.8%; Score 18; DB 2; Length 215;
Best Local Similarity 57.1%; Pred. No. 3.7e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
DB 160 VAISAEF 166

RESULT 96
Q6N9P6 PRELIMINARY; PRT; 216 AA.
AC Q6N9P6;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=RPAL494;
OS Rhodospseudomonas palustris.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Rhodospseudomonas.
OX NCBI_TaxID=1076;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CGA009; ATCC BAA-98;
RX PubMed=14704707; DOI=10.1038/nbt923;
RA Larimer F.W., Chain P., Hauser L., Iamerdin J.E., Malfatti S., Do L.,
RA Land M.L., Pelletier D.A., Beatty J.T., Lang A.S., Tabita F.R.,
RA Gibson J.L., Hanson T.E., Bobst C., Torres y Torres J.L., Peres C.,
RA Harrison F.H., Gibson J., Harwood C.S.;
RT "Complete genome sequence of the metabolically versatile
RT photosynthetic bacterium Rhodospseudomonas palustris.";
RL Nat. Biotechnol. 22:53-61(2004).
DR EMBL; BX572597; CAE26936.1; -.
DR InterPro; IPR009013; Viral_shaft.

```

```

RA Dewhirst F.E., Frazer C.M.;
RT "Complete genome sequence of the oral pathogenic bacterium
RT Porphyromonas gingivalis strain W83.";
RL J. Bacteriol. 185:5591-5601(2003).
DR EMBL; AE017178; AAQ66820.1; -.
DR TIGR; PG1823; -.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 217 AA; 23686 MW; DA473C25FF8851DC CRC64;

Query Match 81.8%; Score 18; DB 2; Length 217;
Best Local Similarity 57.1%; Pred. No. 3.7e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
DB 69 VGAAAEF 75

RESULT 99
Q89ZP3 PRELIMINARY; PRT; 219 AA.
AC Q89ZP3;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein.
GN OrderedLocuNames=BT4333;
OS Bacteroides thetaiotaomicron.
OC Bacteria; Bacteroidetes; Bacteroides (class); Bacteroidales;
OC Bacteroidaceae; Bacteroides.
OX NCBI_TaxID=818;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VPI-5482 / ATCC 29148;
RX MEDLINE=22550859; PubMed=12663928; DOI=10.1126/science.1080029;
RA Xu J., Bjursell M.K., Himrod J., Deng S., Carmichael L.K.,
RA Chiang H.C., Hooper L.V., Gordon J.I.;
RT "A genomic view of the human-Bacteroides thetaiotaomicron symbiosis.";
RL Science 299:2074-2076(2003).
DR EMBL; AE016944; AAQ79438.1; -.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 219 AA; 25376 MW; 5BDCE2094E688E55 CRC64;

Query Match 81.8%; Score 18; DB 2; Length 219;
Best Local Similarity 57.1%; Pred. No. 3.7e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
DB 17 VTVAEEF 23

RESULT 100
Q8FUG5 PRELIMINARY; PRT; 219 AA.
AC Q8FUG5;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN OrderedLocuNames=CE0055;
OS Corynebacterium efficiens.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=152794;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=XS-314;
RX MEDLINE=22723752; PubMed=12840036; DOI=10.1101/gr.1285603;
RA Nishio Y., Nakamura Y., Kawarabayashi Y., Usuda Y., Kimura E.,
RA Sugimoto S., Matsui K., Yamagishi A., Kikuchi H., Ikeo K.,
RA Gojobori T.;
RT "Comparative complete genome sequence analysis of the amino acid

```

```

RT replacements responsible for the thermostability of Corynebacterium
RT efficiens.";
RL Genome Res. 13:1572-1579(2003).
DR EMBL; AP005214; BAC16865.1; -.
DR InterPro; IPR001763; Rhodanese-like.
DR Pfam; PF00581; Rhodanese; 1.
DR SMART; SM00450; RHOD; 1.
DR PROSITE; PS0206; RHODANESE_3; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 219 AA; 23127 MW; BE96467A6760F332 CRC64;

Query Match 81.8%; Score 18; DB 2; Length 219;
Best Local Similarity 57.1%; Pred. No. 3.7e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VXXAAEF 7
DB 53 VRSAAEF 59

```

Search completed: June 13, 2005, 13:59:45
Job time : 122 secs